

**EPIDEMIOLOGY
AND
ENVIRONMENTAL TOBACCO SMOKE**

2503012331

THIS ISSUE BINDER IS INTENDED TO PROVIDE A BASIC,
COMPREHENSIVE REVIEW OF THE SCIENTIFIC LITERATURE
REGARDING A SPECIFIC TOPIC ON ETS AND THE HEALTH OF
NONSMOKERS.

PRIMARY STUDIES AND REVIEWS HAVE BEEN ANNOTATED TO
IDENTIFY (1) USEFUL OR HELPFUL INFORMATION
(UNDERLINING) AND (2) ADVERSE RESULTS OR OPINIONS
(BRACKETING).

2503012332

TABLE OF CONTENTS

2503012333

TABLE OF CONTENTS

TAB

| | | |
|------|---|---|
| I. | STATISTICS AND EPIDEMIOLOGY | 1 |
| | "Statistical Significance and Confidence Intervals" | |
| II. | ETS* | 2 |
| | Scientific Method | |
| | Inadequacies of ETS Studies | |
| | References | |
| III. | ESSAYS ON ETS EPIDEMIOLOGY | 3 |
| | E. Wynder and G. Kabat | |
| | N. Mantel | |
| IV. | GLOSSARY OF TERMS | 4 |
| | Definitions | |
| V. | A DICTIONARY OF EPIDEMIOLOGY | 5 |
| VI. | APPENDIX | 6 |
| | <u>News and Numbers (excerpts)</u> | |

* For analysis and criticism of the epidemiologic studies on ETS and lung cancer, heart disease and respiratory disease in children, see specific Issues Binders.

2503012334

2503012335

STATISTICS AND EPIDEMIOLOGY

"Nevertheless, in a real sense, statistics is the study of populations, or aggregates of individuals, rather than of individuals. Scientific theories which involve the properties of large aggregates of individuals, and not necessarily the properties of the individuals themselves . . . are essentially statistical arguments, and are liable to misinterpretations as soon as the statistical nature of the argument is lost sight of."

Sir Ronald Fisher

2503012336

NOTICE

This material may be
protected by copyright
law (Title 17 U.S. Code).

June 9, 1986 Vol. 144 THE MEDICAL JOURNAL OF AUSTRALIA

618

Statistical significance and confidence intervals

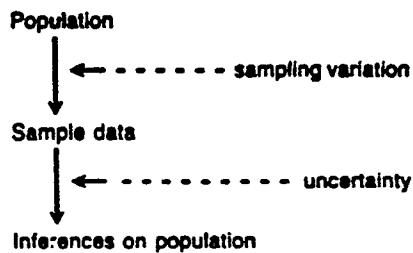
Many papers in the Journal use statistical methods and one of the aims of the review process is to try to ensure that appropriate methods have been used. Often papers report results of comparative studies that are designed to answer questions such as whether one treatment is superior to another for a particular disease, or whether there is an association between some form of behaviour (for example, taking regular exercise or smoking) and the occurrence of some disease. Comparative studies are almost invariably carried out on a sample of individuals who are chosen from the population of individuals to whom it is intended to generalize the results. Data are collected on the sample in order to make inferences on the population. Valid inferences can only be drawn if the sample is chosen in such a way that it is representative of the population. Otherwise a bias could occur; epidemiological methods are designed to eliminate such biases.

Since the aim of a statistical analysis is to make inferences, it is paramount to express whatever inferences that can be drawn in the most informative way. There are several methods of statistical inference, but the two that are most commonly used are significance testing and confidence interval estimation. The former is well known and is featured by quoting *P* values. Many authors appear to be under the impression that a profusion of *P* values is necessary; regrettably this impression has been bolstered in the past by editors of biological journals. Significance testing has its place but, as mentioned by Healy in 1978,¹ "it is widely agreed among statisticians (if less so among the more naive users of statistics) that significance testing is not the be-all and end-all of the subject". In this leading article I would like to discuss the characteristics of both methods of inference, show that a confidence interval contains the result of a significance test, but not vice versa, and suggest that confidence intervals are the answers to the more interesting questions that data can be used to answer.

Any particular study is based on a particular sample; however, it is useful to imagine that the study is repeated with a different sample being selected each time. These hypothetical studies will give different results because they contain different individuals, and individuals vary in any characteristic because of biological variability. The differences are termed sampling variability. It follows then that the results that are obtained from a particular sample can only be taken as an approximation to the actual situation in the whole population. Statistical methods are concerned with assessing the degree of approximation and

what may be reasonably inferred, given that a different sample would have produced a different result.

The methods are based on the assumption that it is a matter of chance which particular subjects are in the sample that is being studied, and the sampling variability is thus random variation which is determined by the laws of probability. Therefore, the inferences are expressed in terms of probability. The situation is illustrated below.



Taking a sample from the population involves sampling variation. As a consequence of this, inferences from the sample data back to the population involve uncertainty.

A statistical analysis may be thought of as asking questions of the data. In an investigation that compares two groups for the mean value of, for example, blood pressure or the prevalence of some disease, three questions may be posed: Is there a difference between the groups?; How large is the difference?; and How accurately is the size of the difference known?.

As expressed, the first question expects the answer "yes" or "no"; although the answer cannot be given in precisely these terms, it is often reduced to two possibilities. The appropriate methodology is the significance test. The second question expects a numerical value to be the answer. This is an estimate and, as it is a single value, is referred to as a point estimate. In effect, the third question asks how reliable this point estimate is; the answer is a range of values which is referred to as an interval estimate or a confidence interval.

These questions represent two approaches to inference: hypothesis testing and estimation. Although at first sight they appear to be quite different, in concept they have much in common. Both make inferential statements about the value of a parameter. (A parameter is an unknown quantity which partly or wholly characterizes a population, for example, a mean or a measure of association.)

The significance test is an appropriate technique when there is an a priori hypothesis to test. For the purpose of the statistical test this hypothesis is expressed in null form — such as when no difference exists between groups — and the test evaluates whether the

data are consistent with the null hypothesis. If the data differ markedly from those which would be expected under the null hypothesis, to the extent that the probability of such an extreme result is low, then it is said that the result is statistically significant. Probability is measured on a continuum between 0 and 1, but in significance testing a probability is considered low if it is less than conventional values such as 0.05 (5%) or 0.01 (1%). A significant result is equated with the rejection of the null hypothesis or the claim of a real effect. By definition, when the null hypothesis is true, significant results will occur by chance with the same relative frequency as the significance probability. That is, real effects will be claimed when the null hypothesis is true; however, the probability of this error (type I) is determined in the data analysis.

One disadvantage of a significance test is that it may fail to detect a real effect; that is, although the null hypothesis is false, the evidence is not strong enough to reject it. The probability of this error (type II) can be controlled at the design stage only, by appropriate selection of the sample size, and may be quite large. Thus, the trap of equating non-significance with no effect must be avoided; failure to reject the null hypothesis is not the same as accepting it.

In the approach of confidence interval estimation no particular hypothesis is considered; rather, the emphasis is on estimating those values of the parameter with which the data are consistent. These values form a range — the confidence interval. The range is calculated so that there is a high probability — conventionally 95% or 99% — that it contains the true value of the parameter.

A significance test is essentially a test of whether the data are consistent with a specified parameter value, and the confidence interval contains those parameter values with which the data are consistent. Therefore, a 5% significance test and a 95% confidence interval contain some information in common: significance implies that the null hypothesis value is outside the confidence interval; non-significance implies that the null hypothesis value is within the confidence interval. However, the confidence interval contains more information because it is equivalent to performing a significance test for all values of the parameter, not just a single value. A confidence interval enables a reader to see how large the effect may be, not simply whether it is different from zero.

The limitations of the interpretations that are provided by a significance test may now be considered.

The difference is significant. This means that there is a difference or, in other words, the size of the difference is not zero. We know no more than this. The difference may

be large and of great importance or it may be small and of no practical importance. It is unsatisfactory that the test provides no way of distinguishing between these quite different possibilities.

The difference is not significant. This means that there is insufficient evidence to enable us to conclude that there is a difference. So the difference may well be zero. But this is not the same as saying that it is zero. The true difference may be quite large. Again, it is unsatisfactory that this possibility is not addressed.

The conclusions that may be drawn from a significance test are considered to be incomplete because it is rarely that one is interested solely in whether a null hypothesis is or is not true; indeed in many cases it may be recognized at the outset that the null hypothesis is unlikely to be true. Rather, the question is how large is the difference and is it possibly large enough to be important? The emphasis is on measuring rather than on testing. The addition of the concept of an important difference to that of a null hypothesis means that there are four possible interpretations to an analysis: (a) the difference is significant and large enough to be of practical importance; (b) the difference is significant but too small to be of practical importance; (c) the difference is not significant but may be large enough to be important; and (d) the difference is not significant and also not large enough to be of practical importance.

The size of difference that is considered to be large enough to be important is a matter for debate, and genuine differences of opinion may arise. It is a medical, not a statistical, question, although a medical statistician who is experienced in the subject area could contribute to setting a value. The fact that agreement on a unique value may be impossible in no way detracts from the argument. In fact, expressing the results as a confidence interval enables interpretations to be made for any particular value that is considered appropriate.

These possibilities are illustrated in the Figure where the confidence intervals are shown. The significant and non-significant cases are distinguished by the confidence intervals that exclude or include zero respectively. The main point is that in each case the confidence interval gives the range of possible values for the true difference. Of particular concern is (c). Here there may be no true difference or there may be a large, important difference. In other words the study is completely inconclusive. Such a possibility is missed by the simple expression "not significant" with its lure of equating this falsely with "no effect". This situation will arise with a study that is carried out on too small a sample and this is why good study design demands attention to sample size to try to prevent the occurrence of an inconclusive result. Altman found that it was common for undue emphasis to be placed on "negative" findings from small studies.²

while Freimen et al. noted that "negative" trials were often too small to constitute a fair test of therapies.³ Similarly, a significance test will contrast (b) as significant and (d) as not significant but fails to recognize that they give essentially the same conclusion — that any difference is too small to be important.

As an example, consider some results which were obtained by Garraway et al. from a clinical trial for the management of acute stroke in the elderly.⁴ Of 155 patients who were managed in a stroke unit, 78 were assessed as independent when they were discharged from the unit compared with 49 of 152 who were managed in a medical unit. The simplest analysis shows that the difference between the success rates of the two units is significant at the 1% level. Therefore, a genuine effect has been established. To appreciate the importance of this effect the advantage of the stroke unit may be measured by the difference between the two units in the percentage of subjects who were discharged as independent: 50.3% - 32.2% = 18.1%. This is the point estimate. The accuracy of this estimate is given by its standard error (5.5) and the 95% confidence limits (7.3% and 28.9%). Thus, the gain could be as large as 29% or as small as 7%.

Recently, Gardner and Altman have argued against the excessive use of hypothesis testing and urged a greater use of confidence intervals.⁵ In an appendix to their paper they give methods to calculate confidence intervals for the commonly occurring two-sample comparisons.

In presenting the main results of a study it is good practice to provide confidence intervals rather than to restrict the analysis to significance tests. Only by so doing can authors give readers sufficient information for a proper conclusion to be drawn; otherwise readers have to rely upon the authors' own interpretation.³ Therefore, intending authors are urged to express their main conclusions in confidence interval form (possibly with the addition of a significance test, although strictly that would provide no extra information). One of the aims of the Journal's statistical review process will be to ensure that where possible this is done.

GEOFFREY BERRY

Associate Professor of Biostatistics
School of Public Health and Tropical Medicine
The University of Sydney

1. Healy MJR. Is statistics a science? *J R Statist Soc A* 1978; 141: 385-393.
2. Altman DG. Statistics in medical journals. *Stat Med* 1982; 1: 59-71.
3. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. *N Engl J Med* 1978; 299: 690-694.
4. Garraway WM, Akhtar AJ, Prentou RJ, Hockey L. Management of acute stroke in the elderly: preliminary results of a controlled trial. *Br Med J* 1980; 280: 1040-1043.
5. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br Med J* 1986; 292: 746-750.

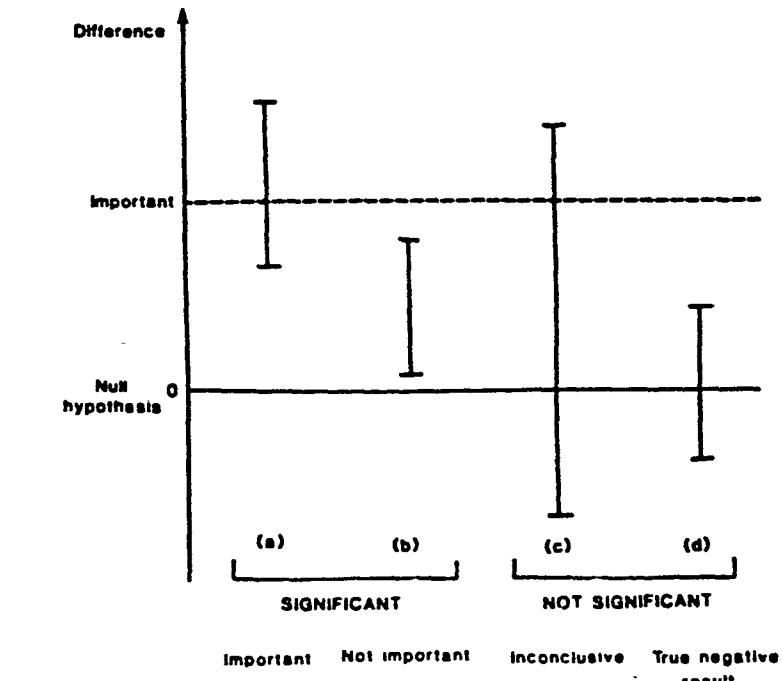


FIGURE: Confidence intervals showing four possible conclusions in terms of statistical significance and practical importance.

2

2503012339

ENVIRONMENTAL TOBACCO SMOKE

Scientific Method

Scientific inquiry within an epidemiologic study begins with framing what is called a "null-hypothesis." The null-hypothesis states, in this instance, that ETS is not associated with a given disease state (e.g., lung cancer, heart disease, etc.). Data are then collected and analyzed in order to test i.e., reject or accept, the hypothesis.

One method which is used to assess the relationship of the collected data to a given hypothesis is the test for statistical significance.* Simply put, if the data examined yield a statistically significant result (here, the relationship between ETS exposure and a disease state), then the scientist is permitted, on the basis of those data, to reject the null-hypothesis. If the statistical test is not significant, then the data do not support rejection of the null hypothesis.

* By convention, a 'p' (probability) value less than 0.05 is deemed statistically significant. A 'p' value less than 0.05 means that the observed results would occur by chance less than 5 times out of 100.

"Confidence limits" are the values between which the risk value can be expected to fall 95% of the time based on the variability of the underlying data. When the 95% confidence limits are both greater and less than 1.00, the risk value is considered not statistically significant, i.e., the results are likely to be due to chance and do not support a judgment regarding an association between exposure and disease.

2503012340

There is no "absolute proof" involved, and there is nothing immutable about the concept of significance testing. Statistical significance is, after all, a convention. But the concept is illustrative, especially in the case of the association between ETS exposure and lung cancer. To date, there have been 28 published reports on ETS and lung cancer, and only five have achieved statistical significance. It is clear that the preponderance of data do not permit rejection of the null hypothesis, i.e., there is no association between ETS exposures and lung cancer. In addition, virtually all of the individual risks reported in such studies are less than 2, which, to the epidemiologist, suggests a "weak" association which is probably the result of bias or confounding of factors unrelated to ETS.

Inadequacies of ETS Studies

Epidemiologic studies are notoriously unreliable in outcome. An observed relative risk of less than 1.5-2.0 (some would up to 3.0) is inadequate to reject the hypothesis of no effect. The overall relative risk calculated across studies is well below a minimal value for seriously attributing it to the presence of a real effect, i.e., it is within the range easily due to the "noise" in epidemiologic data resulting from the limitations and vagaries intrinsic to the methodology and its application. This same conclusion also applies to nearly all of the studies on an individual basis. Another reason for conservative interpretation of the ETS studies is that several studies are of poor quality (good textbook examples of how not to do an epidemiologic study) and some were originally designed for a different, or broader, purpose than assessing health risks from ETS exposure.

2503012341

Sources of bias are present to varying degrees in most of the studies. Lung cancer patients may tend to overstate their exposure to spousal smoking as an explanation for their illness. Bias may result from depending on memory recall of a subject's exposure to spousal smoking. Estimates of relative risk may differ markedly between data collected from the subjects and data obtained from a surrogate, such as their children. Histologic verification of lung cancer was not conducted in all studies and the error rate may be substantial, e.g., 13% of the lung cancer cases in the case-control study of Garfinkel et al. were found to be incorrectly diagnosed when the histology was reviewed by one of the authors. (From: Summary of Public Docket Comments, Draft Risk Assessment, U.S. EPA, Dec. 1990.)

Peter Lee, a statistician and epidemiologist from the United Kingdom, has argued that the increased risks reported in various epidemiologic studies are the result of an inherent bias in study design rather than the result of any genuine effect from exposure to ETS.¹⁻⁵ Lee presents data which indicate that the reported risks cannot be explained on the basis of either ETS exposure or dose for the nonsmoker. It is Lee's contention that the reported "risks" are the result of bias caused by a small number of smokers who are misreported in the studies as nonsmokers.

Other kinds of misclassification may contribute to the reported increase in lung cancer risks among nonsmokers, according to several scientists. For example, none of the studies on ETS and lung cancer provides direct observational information on ETS exposures. Instead, spouses, next-of-kin or friends are asked to estimate the amount of ETS to which they think the subject was exposed. Such estimates may lead to a kind of misclassification,

2503012342

called exposure misclassification,⁶ which has been shown by Garfinkel⁷, Friedman⁸ and others⁹⁻¹² to lead to improper indices of exposure and incorrect estimations of risk. In Garfinkel's study, for example, relative risks varied from 0.83 and 0.77 when the women with lung cancer or the husband was the respondent, to a risk of 3.57 when a son or daughter responded.¹³ That means that the reported risk for lung cancer in the women exposed to ETS was less than for women not exposed when either the women's or their husband's estimates were used.

Dr. S. James Kilpatrick, a biostatistician from the Medical College of Virginia, has analyzed another form of misclassification, called differential misclassification, which results "from the tendency of respondents to inflate the amount of ETS exposure for lung cancer cases and deflate the report of exposure for controls."⁶ Similarly, Dr. Ernst Wynder, President of the American Health Foundation, notes that "relatives of a nonsmoking lung cancer patient are more likely to report passive inhalation exposure on the part of their relative than are relatives of a control patient."¹⁴

A more subtle form of potential bias is known as "publication bias", which stems from the apparent failure by journals to publish studies which report negative or weakly positive results.^{15,16} Scientists have recently expressed concern over the growing trend among such journals to overemphasize (and hence to publish) only those studies which report positive increases in risk.^{17,18} Published studies which are combined for meta-analyses

2503012343

therefore may not truly represent all investigations on the issue of ETS exposures and lung cancer.

Most of the epidemiological studies on ETS and lung cancer have failed to consider age differences, diet, occupation and exposures to indoor or outdoor pollution as potential confounding elements. The importance of such factors is underscored by recently published reports from Japan and China.¹⁹⁻

²⁴ The reports suggest that indoor pollution generated by kerosene heaters, coal stoves, liquified petroleum gas and exposures to cooking oil vapors may be responsible for the increased risk of lung cancer among Oriental women. Moreover, in 1989, researchers in the U.S. reported that nonsmokers living with smokers consumed less carotene (Vitamin A) than did nonsmokers who lived with other nonsmokers. They concluded that "dietary beta-carotene intake is a potential confounder and should be measured whenever possible in studies of the relation between passive smoking and lung cancer."²⁵

Dr. Karl Uberla of Germany recently explained why any attempts to generalize about the significance of reported results of epidemiological studies on ETS and nonsmoker lung cancer will likely remain unconvincing, due to scientific deficiencies in each of the studies.²⁶ He wrote:

The majority of criteria for a causal connection are not fulfilled. There is no consistency, there is a weak association, there is no specificity, the dose-effect relation can be viewed controversially, bias and confounding are not adequately excluded, there is no intervention study, significance is only present under special conditions and the biologic plausibility can be judged controversially.

2503012344

Given these difficulties in interpretation, it is therefore not surprising that an eminent statistician should conclude that "it is unlikely that any epidemiological study has been, or can be, conducted which could permit establishing that the risk of lung cancer has been raised by passive smoking. Whether or not the risk is raised remains to be taken as a matter of faith according to one's choice."¹⁵

Thus, proponents of the ETS health issue are confronted with weak associations and generally statistically nonsignificant risks in epidemiological studies on ETS. They are nevertheless forced to posit a causal mechanism for their theoretical model regarding health risks. They find no support in data from the actual exposure studies on ETS which suggest that an average nonsmoker is exposed, for example, to the nicotine equivalent of one one-hundredth to one one-thousandth (or less) of a single cigarette per hour. Such exposure data suggest that there is no conclusive biological plausibility to the ETS health claim, and that the reported risks in epidemiological studies may be artefactual, and probably due to bias and unconsidered confounders.

2503012345

REFERENCES

1. Lee, P., "Passive Smoking and Lung Cancer: Problems in Interpreting the Epidemiological Data," Presentation to Toxicology Forum, Washington, D.C., February, 1987.
2. Lee, P., "Does Breathing Other People's Tobacco Smoke Cause Lung Cancer?," Br Med J 293: 1503-1504, 1986.
3. Lee, P., "Misclassification as a Factor in Passive Smoking Risk," Lancet II: 867, 1986.
4. Lee, P., "Lung Cancer and Passive Smoking: Association or Artefact Due to Misclassification of Smoking Habits?," Toxicology Letters 35: 157-162, 1987.
5. Lee, P., "Passive Smoking and Lung Cancer Association: A Result of Bias?," Human Toxicol 6: 517-524, 1987.
6. Kilpatrick, S., "Misclassification of Environmental Tobacco Smoke Exposure: Its Potential Influence on Studies of Environmental Tobacco Smoke and Lung Cancer," Toxicology Letters 35: 163-168, 1987.
7. Garfinkel, L., et al., "Involuntary Smoking and Lung Cancer: A Case-Control Study," JNCI 75(3): 463-469, 1985.
8. Friedman, G., et al., "Prevalence and Correlates of Passive Smoking," Am J Public Health 73(4): 401, 1983.
9. Pron, G., et al., "The Reliability of Passive Smoking Histories Reported in a Case-Control Study of Lung Cancer," Am J Epidemiol 127(2): 267-273, 1988.
10. Schenker, M., et al., "Assessment of Environmental Tobacco Smoke Exposure in Epidemiologic Studies," Chest 91(2): 313-314, 1987. Abstract.
11. Lerchen, L. and J. Samet, "An Assessment of the Validity of Questionnaire Responses Provided by a Surviving Spouse," Am J Epidemiol 123(3): 481-489, 1986.
12. Sandler, D. and D. Shore, "Quality of Data on Parents' Smoking and Drinking Provided by Adult Offspring," Am J Epidemiol 124(5): 768-778, 1986.
13. Mantel, N., "What is the Epidemiologic Evidence for a Passive Smoking-Lung Cancer Association?" Indoor Air Quality, H. Kasuga (ed.), Springer-Verlag, (Berlin Heidelberg 1990): 341-347.
14. Wynder, E., "Guidelines to the Epidemiology of Weak Associations," Prev Med 16: 211-212, 1987.

2503012346

15. Mantel, N., "Lung Cancer and Passive Smoking," Br Med J 294: 440, 1987.
16. Lee, P., "Deaths in Canada from Lung Cancer Due to Involuntary Smoking," CMAJ 137: 372-373, 1987.
17. Gordis, L., "Challenges to Epidemiology in the Next Decade," Am J Epidemiol 128(1): 1-9, 1988.
18. Newcombe, R., "Towards a Reduction in Publication Bias," Br Med J 295: 656-659, 1987.
19. Shimizu, H., et al., "A Case-Control Study of Lung Cancer in Nonsmoking Women," Tohoku J Exp Med 154:389-97, 1988.
20. Wu-Williams, A., et al., "Lung Cancer Among Women in North-east China," Brit J Can 62:982-87, 1990.
21. Sobue, T., "Association of Indoor Air Pollution and Passive Smoking With Lung Cancer," Gan no Rinsho 36(3):329-33, 1990. Translation.
22. Du, Y., "Indoor Air Pollution and Woman Lung Cancer," Indoor Air 90 Proceedings I, Toronto, CA.: 59-64, 1990.
23. Mumford, J., et al., "Lung Cancer and Indoor Air Pollution in Xuan Wei, China," Science 235: 217-220, 1987.
24. Gao, Y., et al., "Lung Cancer Among Chinese Women," Int J Cancer 40: 604-609, 1987.
25. Sidney, S., et al. "Daily Intake of Carotene in Nonsmokers with and without Passive Smoking at Home," Am J Epi 129(6): 1305-1309, 1989.
26. Uberla, K., "Epidemiology: Its Scope and Limitations for Indoor Air Quality," Indoor Air Quality: Symposium (Buenos Aires, NAS of Buenos Aires, 1989): 45-60.

2503012347

3

2503012348

ESSAYS ON ETS

10382497

2503012349

H. Kasuga (Ed.)

Indoor Air Quality

With 155 Figures and 190 Tables

Springer-Verlag Berlin Heidelberg New York
London Paris Tokyo Hong Kong

2503012350

Environmental Tobacco Smoke and Lung Cancer: A Critical Assessment*

E. L. Wynder and G. C. Kabat

Summary

The possibility that exposure to environmental tobacco smoke (ETS) may increase the lung cancer risk of nonsmokers has become a cause of public concern. It is unknown whether the levels of carcinogens in the diluted sidestream smoke of tobacco products that reach the nonsmoker's lung are sufficient to induce cancer. Available epidemiologic studies suggest a slight increase in the relative risk of lung cancer in nonsmokers due to exposure to ETS created by a smoking spouse. However, not all studies have found a significant association. The epidemiologic studies are examined in the light of the criteria of judgment of causality, including strength of association, consistency, temporality, methodological issues, and biological plausibility. Suggestions for further research, including studies in high-exposure populations and greater attention to histology, are proposed.

Introduction

Epidemiologists, chemists, biologists, physiologists, physicians, and public health officials have given much attention to the association of environmental tobacco smoke (ETS) exposure and the development of lung cancer in nonsmokers. A biological basis for such an association clearly exists because smoke constituents demonstrated to be carcinogenic in laboratory animals are inhaled and retained by the nonsmoker. Metabolites of tobacco-specific smoke constituents have been identified in the saliva, blood, and urine of nonsmokers after exposure to ETS (Greenberg et al. 1984; Hoffmann et al. 1984; National Academy of Sciences 1986; USDHHS 1987; Sepkovic et al. 1988). Several epidemiological studies have found a positive association between ETS exposure – usually defined as being due to a smoking spouse – and lung cancer (Hirayama 1981; Trichopoulos et al. 1981; Correa et al. 1983; Sandler et al. 1985; Garfinkel et al. 1985; Akiba et al. 1986; Dalager et al. 1986; Pershagen et al. 1987). Other studies have found no significant association (Garfinkel 1981; Chan and Fung 1982; Koo et al. 1983; Kabat and Wynder 1984; Wu et al. 1985; Lee et al. 1986). No consistent association has been reported for lung cancer and exposure to ETS in childhood, which might be expected to exert a greater effect, especially when followed by exposure throughout adulthood. Of course, recall of ETS exposure in childhood is more difficult than recall of such exposure in adulthood.

* Research described herein was performed under USPHS, National Cancer Institute Program Project Grant CA-32617.

The epidemiological study of weak associations is burdened with problems that may yield artifactual positive findings or may show negative findings where a real association exists. The association of ETS and lung cancer risk, even if weak, would still be of concern as a public health problem in that most people are at one time or another exposed to smoke from burning tobacco products and the exhaled pollutants of tobacco smokers. A weak association in epidemiology requires careful examination and an understanding of the variables in question and all of the factors influencing the association (Wynder 1987).

In this overview we critically examine the published studies on ETS exposure and lung cancer to determine whether the evidence presented to date permits a sound conclusion as to causation.

General Exposure to ETS

At the outset we need to emphasize that an association between ETS and lung cancer must be deemed possible. A recent survey of self-reported exposure in a hospitalized population revealed that 66% of men and 60% of women had ETS exposure in childhood; 32% of the men and 61% of the women reported ETS exposure in the home in adulthood; and 60% of the men and 62% of the women who worked outside the home reported ETS exposure at work (Kabat and Wynder, unpublished data, 1987).

Critical Assessment

The first Surgeon-General's Report on Smoking and Health, published in 1964 (USPHS 1964), clearly delineated the criteria of judgment for causality. These criteria included: the magnitude of the association, consistency, temporality, and biological plausibility. Since these criteria were considered necessary to prove causation for a strong association, namely, active smoking and lung cancer, they should be equally required to determine the causality of weak associations (Wynder 1987). Let us examine the epidemiological evidence linking ETS with lung cancer in respect to these criteria.

Strength of the Association

An association is generally considered weak if the odds ratio is under 3.0 and particularly when it is under 2.0, as is the case in the relationship of ETS and lung cancer (Table 1). If the observed relative risk is small, it is important to determine whether the effect could be due to biased selection of subjects, confounding, biased reporting, or anomalies of particular subgroups.

Consistency

If an association is real, internal consistency should be apparent within and between different studies. The majority, but not all of the studies of ETS and lung cancer have shown a positive association for ETS-exposure due to a smoking spouse (Table 1). In most of the studies, the confidence interval includes 1.0. While the prospective study by Hirayama (1981a) among Japanese women showed a significant association with the husband's smoking (largely adenocarcinomas), the prospective study among American

2503012352

Table 1. Summary of results of studies relating lung cancer risk in married women to their husbands' smoking habits

| | Relative risk | 95% Confidence interval |
|-----------------------------|---------------|-------------------------|
| <i>Prospective studies</i> | | |
| Hirayama (1981) | 1.63 | 1.25-2.11 |
| Garfinkel (1981) | 1.18 | 0.90-1.54 |
| <i>Case-control studies</i> | | |
| Trichopoulos et al. (1981) | 2.1 | 1.18-3.78 |
| Chan & Fung (1982) | 0.75 | 0.44-1.30 |
| Correa et al. (1983) | 2.03 | 0.83-5.03 |
| Koo et al. (1983) | 1.54 | 0.90-2.64 |
| Kabat & Wynder (1984) | 0.79 | 0.26-2.43 |
| Wu et al. (1985) | 1.2 | 0.6-2.5 |
| Garfinkel et al. (1985) | 1.12 | 0.74-1.69 |
| Lee et al. (1985) | 1.03 | 0.41-2.47 |
| Akiba et al. (1986) | 1.48 | 0.88-2.50 |
| Pershagen et al. (1987) | 1.28 | 0.75-2.16 |

Table 2. Distribution of lung cancer by histologic groups in smokers and never-smokers. (From Kabat and Wynder 1984)

| | Smokers | | Never-smokers | |
|---------------------------------------|----------------------------|-----------------------------|--------------------------|----------------------------|
| | Males (N = 1882) [%] | Females (N = 652) [%] | Males (N = 37) [%] | Females (N = 97) [%] |
| Kreyberg I | 63 | 52 | 35 | 21 |
| Kreyberg II | 32 | 43 | 54 | 74 |
| Mixed and undifferentiated/anaplastic | 5 | 5 | 11 | 5 |

women by Garfinkel (1981) did not. It has been suggested that Japanese and American women are exposed to different levels of ETS due to different conditions in the two countries. Such differences could account for this disparity (Hirayama 1981b).

Within those studies presenting specific histologic analysis, differences exist in respect to the type of lung cancer involved. In active smokers, tobacco smoke exposure has a causative effect predominantly on squamous and small cell types of lung cancer (Kreyberg I), with a lesser, though still significant causative effect on the glandular type (Kreyberg II) (Wynder and Stellman 1977). Among nonsmokers, however, the glandular type of lung cancer predominates among both men and women (Kabat and Wynder 1984) (Table 2). The effect of ETS would thus be expected to be primarily responsible for the higher rate of adenocarcinomas among nonsmokers. The studies by Dalager et al. (1986) and Pershagen et al. (1987), however, suggest that the effect of ETS exposure is limited to induction of squamous cell lung cancer (Table 3). If this were, in fact, the case, then only the squamous or small cell type of lung cancer in nonsmokers

2503012353

Table 3. Histology-specific odds ratios for spouse smoking from two studies

| Study | Histologic type | N | Odds ratio | 95% C.I. |
|----------------------------|----------------------------|----|------------|------------|
| Dalager et al. (1986) | Adenocarcinoma | 16 | 1.02* | 0.33- 3.16 |
| | Squamous & Small Cell Ca. | 14 | 2.88* | 0.91- 9.10 |
| | Other | 18 | 1.31* | 0.48- 3.57 |
| Pershagen et al. (1987) | Squamous or Small Cell Ca. | 20 | 3.3 | 1.1 -11.4 |
| | Other | 47 | 0.8 | 0.4 - 1.5 |

would be affected by ETS. Clearly, it is important that investigations of the effect of ETS exposure on lung cancer development in nonsmokers take histology into account, so as to determine whether an effect of ETS is limited to certain histological types.

Since smoking is more prevalent in lower income groups, at least among men, lung cancer in nonsmoking women in these groups should have a higher incidence. Thus, the influence of the level of education on smoking habits in the examined population needs to be considered as a possible confounder. Few studies to date have done this.

Methodological Issues

A particular concern in weak associations is reporting bias, that is, potentially differential reporting of exposures between cases and controls. In terms of ETS, does the lung cancer patient report exposure to tobacco smoke, be it at work, at home, at social functions, in childhood or adulthood, differently than the control? The case is likely to have a different attitude toward this question than does the control, a handicap not applicable to prospective studies. It needs to be determined whether the case's attitude towards questions on ETS exposure leads to under- or overreporting. Cases are likely to underreport their own smoking (Lee 1987), and they may tend to overreport their exposure to ETS and other potential hazards that could account for their illness. In studies that use proxy reports, different relatives may respond differently. Garfinkel et al. (1985) provides some insight into this phenomenon by showing that if the response came from the patient, the odds ratio was 1.0, if from the husband it was 0.92, and if from the daughter or son, 3.19 (Table 4). More work is needed on the validity of ETS-exposure information obtained from different relatives before we can evaluate which of these relative risks is closer to the truth.

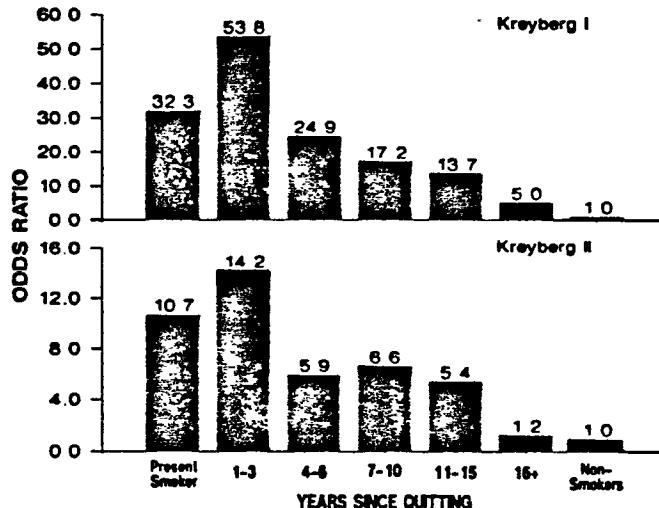
In general, possible reporting bias represents a serious problem in case-control studies because it can produce a systematic artefact. It is particularly worrisome in that it cannot be effectively measured.

We also need to consider misclassification that can occur in both retrospective and prospective studies. Lee has proposed (Lee et al. 1986; Lee 1987) that the reported ETS effect on lung cancer risk can be explained by a misclassification of smokers as nonsmokers. According to these studies, a substantial percentage of respondents misrepresent their smoking habits. Using a 10.0% misclassification rate of ex-smokers as self-reported never smokers coupled with the concordance of spouses' smoking habits,

2503012354

Table 4. Data from Garfinkel et al. (1985) by type of respondent

| Husband's smoking habits at home | | | |
|----------------------------------|------------|------|------------|
| | N of cases | OR | 95% C.I. |
| Self | 16 | 1.00 | 0.55- 1.74 |
| Husband | 34 | 0.92 | 0.63- 1.34 |
| Daughter/son | 48 | 3.19 | 0.91-11.19 |
| Other | 36 | 0.77 | 0.57- 1.03 |

Fig. 1. Odds ratio of male ex-smokers for Kreyberg I ($N = 687$) and Kreyberg II ($N = 301$) lung cancer by years since quitting (controls = 6534). Source: American Health Foundation data

Lee calculated that an apparent increase in lung cancer risk can be obtained among nonsmokers married to smokers that approximates the increased risk observed in a number of epidemiologic studies (Lee 1987). At the extreme, Garfinkel et al. (1985) showed that 40% of lung cancer cases classified as "nonsmokers" in the hospital chart were in fact smokers as determined by interview. Although such a high rate of misclassification does not occur when cases are interviewed personally, to some extent denial is likely to occur even then, particularly among ex-smokers who had stopped smoking ten or more years ago. The risk of lung cancer among long-term ex-smokers, and even among ex-smokers who quit more than 16 years earlier, does remain elevated above the rate among those who never smoked (Fig. 1). Denial of past smoking may also not be uncommon in populations where smoking is or was socially unacceptable, as is the case among older Japanese women.

2503012355

Table 5. Percent of lung cancer cases who never smoked by histologic group (A.H.F. data)

| | Males | | | | Females | | | |
|-----------|-------|-----|-------|-----|---------|-----|-------|-----|
| | KI* | | KII** | | KI* | | KII** | |
| | [%] | N | [%] | N | [%] | N | [%] | N |
| 1969-1973 | 1.2 | 488 | 5.6 | 142 | 10.7 | 103 | 23.7 | 76 |
| 1974-1976 | 1.6 | 887 | 3.0 | 305 | 16.4 | 263 | 25.3 | 146 |
| 1977-1980 | 2.1 | 628 | 4.6 | 390 | 5.6 | 231 | 22.0 | 245 |
| 1981-1985 | 1.4 | 725 | 5.6 | 463 | 6.8 | 311 | 16.6 | 284 |

* Kreyberg I

** Kreyberg II

Another problem for epidemiologists involves subgroup analysis (Stallones 1987). Investigators are likely to examine numerous subgroups, and then prefer to present those subgroups that best fit the hypothesis. This tendency represents an inherent problem in epidemiology. The investigator should at a minimum give an idea of how many subgroups were originally examined and how many subgroups were discarded.

Temporality

One of the factors that led to the conclusion that active smoking causes lung cancer was that the increase in cigarette consumption preceded the increase in lung cancer rates, first in men and later in women. Enstrom (1979) has reported an increase in the lung cancer rate in nonsmokers over recent years, suggesting that factors in addition to personal cigarette smoking influence lung cancer mortality rates. The groups examined, however, are not strictly comparable, and misclassification of smokers as nonsmokers in the national surveys needs to be considered. Our data from a long-term, hospital-based case-control study do not indicate an increase in the percentage of male nonsmokers with lung cancer in either of the two main histologic groupings (Kreyberg I and II) over the last 30 years (Table 5).

In fact, the percentage of nonsmokers with lung cancer among women has declined, which may be a consequence of the diminishing pool of women who have never smoked.

Biological Plausibility

Several studies have demonstrated that most tumorigenic agents are present in undiluted sidestream smoke in higher concentrations than in mainstream smoke (Hoffmann et al. 1983; National Academy of Sciences 1986; Hoffmann and Wynder 1986) (Table 6). Biochemical studies indicate that nonsmokers exposed to ETS have levels of nicotine or cotinine in the blood or urine that are about 1/100th the level seen in active smokers (Table 7) (Jarvis et al. 1984; National Academy of Sciences 1986). Some of the nicotine measured in the blood and urine represents nicotine that is absorbed by the saliva of nonsmokers and does not reach the lung directly (Jarczyk et al. 1987). It is important to

2503012356

Table 6. Distribution of compounds in undiluted cigarette mainstream smoke (MS) and sidestream smoke (SS)

Nonfilter cigarettes

| | MS | SS/MS |
|------------------------------|----------------|------------|
| <i>(A) Vapor phase</i> | | |
| Carbon monoxide | 10 - 23 mg | 2.5 - 4.7 |
| Carbon dioxide | 20 - 40 mg | 8 - 11 |
| Benzene | 20 - 50 µg | 10 |
| Formaldehyde | 5 - 100 µg | 0.1 - ~50 |
| Acrolein | 50 - 100 µg | 8 - 15 |
| Acetone | 100 - 250 µg | 2 - 5 |
| Hydrogen cyanide | 400 - 500 µg | 0.1 - 0.25 |
| Hydrazine | 24 - 43 ng | 3.0 |
| Ammonia | 50 - 170 µg | 40 - 170 |
| Methylamine | 11.5 - 28.7 µg | 4.2 - 6.4 |
| Nitrogen oxides | 50 - 600 µg | 4 - 10 |
| N-nitrosodimethylamine | 10 - 180 ng | 20 - 100 |
| N-nitrosopyrrolidine | 2 - 110 ng | 6 - 30 |
| <i>(B) Particulate phase</i> | | |
| Particulate matter | 15 - 40 mg | 1.3 - 1.9 |
| Nicotine | 1 - 2.5 mg | 2.6 - 3.3 |
| Phenol | 60 - 140 µg | 1.6 - 3.0 |
| Catechol | 100 - 350 µg | 0.6 - 0.9 |
| Hydroquinone | 110 - 300 µg | 0.7 - 0.9 |
| Aniline | 360 ng | 30 |
| 2-Toluidine | 30 - 160 ng | 19 |
| 2-Naphthylamine | 4.3 - 27 ng | 30 |
| 4-Aminobiphenyl | 2.4 - 4.6 ng | 31 |
| Benz(a)anthracene | 40 - 70 ng | 2 - 4 |
| Benzo(a)pyrene | 10 - 40 ng | 2.5 - 3.5 |
| N'-Nitrosonornicotine | 120 - 3,700 ng | 0.5 - 3 |
| NNK | 120 - 950 ng | 1 - 4 |
| Cadmium | 100 ng | 7.2 |
| Nickel | 20 - 3,000 ng | 13 - 30 |
| Polonium-210 | 0.03 - 1.0 pCi | ? |

note that nicotine occurs in ETS primarily as a vapor phase constituent rather than in the particulate matter of the aerosol as is the case in mainstream cigarette smoke (Eudy et al. 1987). Measurement of nicotine or its metabolites will, therefore, not reflect the proportional uptake of particulate matter from ETS. In the light of our present knowledge of dose-response in carcinogenesis and because the carcinogenic activity of tobacco smoke as measured in animal systems is relatively low, the question needs to be raised whether the carcinogenic potential of inhaled ETS suffices to induce lung cancer. Hoffmann and Hecht (1985) have proposed nicotine-derived nitrosamines in ETS as organ-specific carcinogens for the lung. It is possible that these chemicals reach the lungs in sufficient dose to induce neoplastic changes. These carcinogens may also be formed endogenously from inhaled or ingested nicotine and appropriate nitrosating agents (Hoffmann and Hecht 1985). Tumor promoters are less likely to play a role in ETS

2503012357

Table 7. Approximate relations of nicotine as a parameter between non-smokers, passive smokers and active smokers*. (From Jarvis et al. 1984)

| Nicotine/cotinine | Non-smokers without ETS exposure No. = 46 | | Non-smokers with ETS exposure No. = 54 | | Active smokers No. = 94 |
|-------------------------|---|---------------------------------|--|---------------------------------|-------------------------------|
| | Mean value | % of active smokers value | Mean value | % of active smokers value | Mean value |
| <i>Nicotine (ng/ml)</i> | | | | | |
| in plasma | 1.0 | 7 | 0.8 | 5.5 | 14.8 |
| in saliva | 3.8 | 0.6 | 5.5 | 0.8 | 673 |
| in urine | 3.9 | 0.2 | 12.1* | 0.7 | 1,750 |
| <i>Cotinine (ng/ml)</i> | | | | | |
| in plasma | 0.8 | 0.3 | 2.0* | 0.7 | 275 |
| in saliva | 0.7 | 0.2 | 2.5** | 0.8 | 310 |
| in urine | 1.6 | 0.1 | 7.7** | 0.6 | 1,390 |

* Differences between non-smokers exposed to ETS compared with non-smokers without exposure

* p < 0.01

** p < 0.001

carcinogenesis than in active smoking because of their much lower concentration. In general, tumor promoters are effective only when applied repeatedly in relatively large amounts.

In considering the existing data on ETS exposure and lung cancer, it is noteworthy that Auerbach et al. (1961) showed only minor histological changes in the bronchial epithelium of nonsmokers and found that the ciliated columnar epithelium that covers their bronchi were largely intact. Deposition of carcinogenic smoke particulates can take place only upon inhibition of the protective functioning of the lung clearance system. Squamous cell lung cancer can arise only from ciliated columnar cells that have undergone squamous metaplasia.

An active smoker with each puff from a cigarette inhales a volume of 35–50 ml of a concentrated aerosol containing 3–5 billion particles per ml that adversely affect the protective cilia and mucous defense system of the bronchi (Ferin et al. 1965). The passive smoker is at no time exposed with such force to such a highly polluted inhalant. Furthermore, ETS particles are more likely to be deposited in the upper respiratory tract and not predominantly in the bronchi as is the case in active smoking. Thus, our respiratory defense system may be able to deal more readily with the relatively lighter deposition of particles and exposure to volatiles in ETS, as the observation by Auerbach et al. (1961) would suggest.

Future Studies

Future epidemiological studies on the association of ETS with lung cancer should attempt to avoid the pitfalls discussed above. The definitive evidence that a factor causes

2503012358

human cancer requires support from descriptive, metabolic, and molecular epidemiology.

Beyond extension of prospective studies, such as those now in progress by Garfinkel and Stellman at the American Cancer Society, we suggest:

- 1) Continuing ongoing case-control studies with special reference to histologic type and careful consideration of methodological issues.
- 2) Estimating the relative importance of ETS exposure in different settings - in the home, in the workplace, in social situations, and during transportation.
- 3) Further studying lung cancer rates among pipe and cigar smokers, and, if feasible, among nonsmokers exposed to ETS from these products.
- 4) Studying lung cancer incidence in groups occupationally exposed to high levels of ETS at their worksite such as waiters, bartenders, train conductors, airplane personnel, and office workers.
- 5) Studying bronchial epithelium in autopsy material of established never-smokers whose exposure to ETS is known.
- 6) Determining the incidence of lung cancer by histological type in confirmed never-smokers.
- 7) Comparing the presence of adducts of tobacco-specific carcinogens with DNA in smokers, passive smokers, and "never-smokers" (Hoffmann and Hecht 1985; Hecht et al. 1987).

In summary, verification of the possible association of ETS and lung cancer represents an important challenge to epidemiologists, laboratory scientists, and public health authorities. The public is entitled to inhale the cleanest possible air regardless of whether ETS is proven to be cancer-inducing. Additional efforts on the part of epidemiologists are required to firmly establish the nature and significance of the reported associations between passive smoking and lung cancer.

References

Akiba S, Kato S, Blot WJ (1986) Passive smoking and lung cancer among Japanese women. *Cancer Res* 46:4804-4807

Auerbach O, Stour AP, Hammond EC, et al (1961) Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. *N Engl J Med* 265:253-267

Chan WC, Fung SC (1982) Lung cancer in non-smokers in Hong Kong. In: Grundman E (ed) *Cancer campaign*, vol 6. *Cancer epidemiology*. Fischer, Stuttgart, pp 199-202

Correa P, Fontham E, Pickle L, Lin Y, Haenszel W (1983) Passive smoking and lung cancer. *Lancet* 2:595-597

Dalager NA, Pickle LW, Mason TJ, et al (1986) The relation of passive smoking to lung cancer. *Cancer Res* 46:4808-4811

Enstrom JE (1979) Rising lung cancer mortality among nonsmokers. *J Natl Cancer Inst* 62:755-760

Eudy LW, Thorne FA, Heavner DL, Green CR, Ingebretsen BJ (1985) Studies on the vapor-phase distribution of environmental nicotine by selected trapping and detection methods. *Pres. 39th Tobacco Chemists Res Conf*, 1985, p 25

Ferin J, Urbankova G, Vlckova A (1965) Influence of tobacco smoke on the elimination of particles from the lungs. *Nature* 206:515-516

Garfinkel L (1981) Time trends in lung cancer mortality among non-smokers and a note on passive smoking. *J Natl Cancer Inst* 66:1061-1066

2503012359

Garfinkel L, Auerbach O, Joubert L (1985) Involuntary smoking and lung cancer: A case-control study. *J Natl Cancer Inst* 75:463-469

Greenberg RA, Haley NJ, Etzel RA, Loda FA (1984) Measuring the exposure of infants to tobacco smoke: nicotine and cotinine in urine and saliva. *N Engl J Med* 310:1075-1078

Hecht SS, Carmella SG, Trushim N, Foiles PG, Lin D, Rubin YM, Chung FL (1987) Investigations on the molecular dosimetry of tobacco-specific N-nitrosamines. International Agency for Research on Cancer, Lyon [IARC Sci Publ 84:423-429]

Hirayama T (1981a) Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study in Japan. *Br Med J* 282:183-185

Hirayama T (1981b) Nonsmoking wives of smokers have a higher risk of lung cancer [Letter]. *Br Med J* 283:916-917

Hoffmann D, Hecht S (1985) Nicotine-derived N-nitrosamines and tobacco-related cancer: Current status and future directions. *Cancer Res* 45:935-944

Hoffmann D, Wynder EL (1986) Chemical constituents and bioactivity of tobacco smoke. In: Zaridze DG, Peto R (eds) *Tobacco: A major international health hazard*. International Agency for Research on Cancer, Lyon, IARC Sci Publ 74:145-65

Hoffmann D, Haley NJ, Brunnemann KD, Adams JD, Wynder EL (1983) Cigarette sidestream smoke: formation, analysis and model studies on the uptake by nonsmokers. US-Japan meeting "New Etiology of Lung Cancer", Honolulu, Hawaii, March 21-23, 1983

Hoffmann D, Haley NJ, Adams JD, Brunnemann KD (1984) Tobacco sidestream smoke: uptake by nonsmokers. *Prev Med* 13:608-613

Jarczyk L, Slerer G, Maltzan C, Luu HT, Adikofler F (1987) Intake of nicotine from ETS via different inhalation routes. Proceedings of Third International Conference on Indoor Air Quality (in press)

Jarvis M, Tunstall-Pedoe H, Feyerabend C, Vesey C, Saloojee Y (1984) Biochemical markers of smoke absorption and self-reported exposure to passive smoking. *J Epidemiol Comm Health* 38:335-339

Kabat GC, Wynder EL (1984) Lung cancer in nonsmokers. *Cancer* 53:1214-1221

Koo LC, Ho JH-C, Saw D (1983) Active and passive smoking among female lung cancer patients and controls in Hong Kong. *J Exp Clin Cancer Res* 4:365-375

Lee PN (1987) Lung cancer and passive smoking: association an artefact due to misclassification of smoking habits? *Toxicol Lett* 35:157-162

Lee PN, Chamberlain J, Alderson MR (1986) Relationship of passive smoking to risk of lung cancer and other smoking-associated diseases. *Br J Cancer* 54:97-105

National Academy of Sciences, National Research Council (1986) *Environmental tobacco smoke. Measuring exposures and assessing health effects*. National Academy Press, Washington DC

Pershagen G, Hrubec Z, Svensson C (1987) Passive smoking and lung cancer. *Am J Epidemiol* 125:17-24

Sandler DP, Everson RB, Wilcox AJ (1985) Passive smoking in adulthood and cancer risk. *Am J Epidemiol* 121:37-48

Sepkovic DW, Axelrad CM, Colosimo SG, Haley NJ (1988) Measuring tobacco smoke exposure: clinical applications and passive smoking. 80th Annual Meeting and Exhibition of the Air Pollution Control Association, New York, NY, June 21-26, 1987 (in press)

Stallones RA (1987) The use and abuse of subgroup analysis in epidemiological research. *Prev Med* 16:183-194

Trichopoulos D, Kalandidi A, Sparros L, MacMahon B (1981) Lung cancer and passive smoking. *Int J Cancer* 27:1

[USDHHS] US Dept Health and Human Services, Public Health Service, Office of Smoking and Health (1984) *The health consequences of smoking: cardiovascular disease. A Report of the Surgeon General*. USDHHS, US Govt. Printing Office, [DHHS Publ No (PHS) 84-50204] Washington DC

[USDHHS] US Dept Health and Human Services, Public Health Service, Centers for Disease Control (1987) *The Health Consequences of Involuntary Smoking, a Report of the Surgeon General*. USDHHS, US Govt. Printing Office, [DHHS (CDC) Publ No 87-8398], Washington DC

2503012360

[USPHS] US Public Health Service. Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service (1964) US Dept of Health, Education, and Welfare, Public Health Service, Center for Disease Control [PHS Publ No 1103] US Govt Printing Office, Washington DC

Wu AH, Henderson BE, Pike MC, Yu MC (1985) Smoking and other risk factors for lung cancer in women. *J Natl Cancer Inst* 74:747-751

Wynder EL (ed) (1987) Workshop on guidelines to the epidemiology of weak associations. *Prev Med* 16:139-212

Wynder EL, Stellman ST (1977) Comparative epidemiology of tobacco-related cancers. *Cancer Res* 37:4608-4622

2503012361

H. Kasuga (Ed.)

Indoor Air Quality

With 155 Figures and 190 Tables

Springer-Verlag Berlin Heidelberg New York
London Paris Tokyo Hong Kong

2503012362

What Is the Epidemiologic Evidence for a Passive Smoking-Lung Cancer Association?

N. Mantel

Summary

Two survey articles of reports on the association of passive smoking with lung cancer have recently appeared, and also a comprehensive report on the subject of environmental tobacco smoke by a committee of the National Research Council of the United States. The observed excess over a relative risk of unity cannot be explained by chance. Nor can it be fully accounted for by a particular source of bias, the false claims of being non-smokers by individuals who were active or ex-smokers. That possible source of bias leads, in one summary survey, to reducing a relative risk of 1.35 to 1.30, but from 1.34 to 1.15 in the National Research Council report. The latter report suggests that statistical significance would no longer obtain, perhaps, particularly, because of other possible biases. However, to get an estimate of the correct relative risk due to passive smoking, allowance has to be made for actual exposure to passive smoking of those not exposed at home. Thus, the 1.30 is adjusted upwards, by 18 in one survey, to 1.53, but by only 8% in the National Research Council report to 1.24. The National Research Council report had given an anticipated relative risk of 1.1 based on dosimetric considerations. But it is suggested here that that could be as low as 1.05, too low to be detected in an epidemiologic investigation - in any case it would be based on hypothetical assumptions.

In November of 1986 there were two near-simultaneous review articles addressing the subject of passive smoking and lung cancer. One was an invited guest editorial by Blot and Fraumeni in the Journal of the National Cancer Institute, the other a contemporary theme discussion by Wald et al. in the British Medical Journal [1, 2].

There was substantial overlapping in the two articles of the various publications on the subject, and on the basis of which the conclusion of a significant positive association was made. The article by Wald et al. gave, perhaps, more statistical detail about the results of the several studies covered. But, to my mind, there was uncritical acceptance of the results of all the studies. Blot and Fraumeni did suggest that there were some flaws in a particular study, that by Hirayama [3], but decided that any inherent biases in that investigation could not have given rise to the observed elevated risk.

From their overall evaluation of 10 case-control studies (all 10 gave results for females, five separately for males as well) and three prospective studies (two of these covered males separately), which provided 20 separate relative risk (actually odds ratio) values, Wald et al. came up with a summary relative risk of lung cancer due to passive smoking of 1.35 (95% limits 1.19 to 1.54). They trim this down to 1.30 on the basis that some of the presumed non-smokers exposed to passive smoking were actually smokers. Then, on the added basis that even those unexposed to passive smoking at home may still have been exposed when away from home, they raise their estimate of relative risk to 1.53. But note that this last modification presupposes the answer, that passive smoking does

elevate the risk. For if it did not, there would be no basis for adjusting the 1.30 or 1.35 upwards to 1.53.

Blot and Fraumeni come up with a similar summary measure of relative risk for passive smoking of 1.3 (95% limits of 1.1-1.5), but elevated to 1.7 (95% limits of 1.4-2.1) for heavy passive smoking. These authors suggest that heavy passive smoking is equivalent, at least in terms of nicotine received, to smoking between 1/2 and 3 cigarettes daily, and estimate that smoking a few cigarettes daily would give rise to a relative risk of about 1.5-fold to twofold.

While Blot and Fraumeni do not address the question of correct reporting of non-smoking status, Wald et al. do, having used this as a basis for lowering the relative risk estimate from 1.35 to 1.30. Based on reports and communications from others, Wald et al. estimate that persons reporting themselves as never having smoked (lifelong non-smokers) comprise 2.1% active smokers plus 4.9% former smokers, for a total of 7% ever smokers among the self-claimed never smokers. Wald et al. estimate that these 7% have a combined relative risk of 2, making the assumption in doing this that the active smokers among the 7% smoked on average only a quarter as much as active smokers generally. The relative risk of 2 for the 7% is computed as a weighted average of 3 for active smokers, 1.5 for former smokers, among the 7%.

If 7% of reported never-smokers were actually ex-smokers or active smokers, which were they - the spouses, say, of smokers or the spouses of non-smokers? In my own critique of Hirayama, I had suggested that this false reporting of non-smoking status would preferentially be among those with smoking spouses [4]. If, for example, the 7% overall misreporting of non-smoking status concentrated among spouses of smokers, it would be somewhat higher among persons with smoking spouses who, nevertheless, claimed to be never smokers. Suppose we take it at 20%, in which case the reported lifelong non-smokers relative risk would be 1.20. It could be substantially higher but for the assumption by Wald et al. that the active smokers among the reported never smokers had sharply reduced levels of smoking. However, Wald et al. were ready to make only a small reduction in relative risk for this factor, from 1.35 to 1.30. Their speculative increase, which might have no basis at all, was much greater, from 1.30 to 1.53.

The effect of false reporting of smoking status, specifically of non-smoking, could be much sharper than what Wald et al. have suggested. In a study of biochemical markers of smoke absorption, Jarvis et al. branded as "deceivers" 21 individuals who claimed to be non-smokers [5]. These 21 displayed biochemical patterns very similar to those of actual smokers, not at all like those of accepted non-smokers. The 100 accepted non-smokers comprised 46 without passive smoking, 54 with. Those 21 would constitute 21/121 or about 17% of the total, and these would be active smokers, not just former smokers, or eightfold greater than the 2.1% Wald et al. postulated. Perhaps in the epidemiologic investigations made, false reporting of non-smoking status is at a much lower level, but it would not take much false reporting to account fully for the seeming association between passive smoking and lung cancer.

Recently, a colleague expressed to me the thought that if passive smoking played no role in lung cancer, why are we not finding many negative associations, nor any significantly negative associations? Actually, six of the 20 relative risks reported in Wald et al. are at 1.00 or smaller. And some of those reported as in excess of 1.00 conceal rates of under 1.00. Thus, relative to the rate shown of 1.23 for the study reported by Garfinkel et al., I have brought out in my own critique that that represented a composite of data for various classes of respondents [6, 7]. Where the woman with lung cancer was herself the respondent (as to her husband's level of smoking) the relative risk was 0.83. Using the husbands' responses, the relative risk was 0.77. It was only on the basis of responses by

2503012364

the sons and daughters, at a time long past when they would have left home, that a relative risk of 3.57 emerged, sufficiently high to raise the overall estimate of relative risk to 1.23. As I indicated in my critique, the replies by the children were more accusatory in nature than revealing of any true relationship.

But even so, it would take 40 large studies to get on average a single seemingly significant negative association of lung cancer with passive smoking, assuming statistical testing at the 5%, two-tailed, level. But we have only 20 evaluations, with many so small that they could not possibly yield any apparently significant protective effect, not even in the unrealistic situation that passive smoking was 100% protective. Suppose a study had a null expectation of only 2 or 3 passive smokers with lung cancer - then there would be some observed number, 5 or 6 or 7 or 8 or more which would be significantly in excess of expectation. But there would be no number, however small or even zero, which would be significantly below expectation. Yet just such low expectations characterize several of the studies reported on by Wald et al. In one study, a relative risk of 2.29 is shown based on only 2 actual cases of lung cancer in passive smokers, expectation 1.20. Another relative risk of 2.45 is based on 3 observed, 1.77 expected. For one prospective study, 4 observed cases have given rise to an estimated relative risk of 3.25, and in another 7 observed cases gave rise to a relative risk of 2.25, suggestive of an expectation little in excess of 3. On the other hand, the four reported risks of under 1.00 had expectations variously of 37.67, 34.08, 6.64 and 13.77.

Of concern to Wald et al. was whether the various relative risks were homogeneous. On this point they cite a chi-square test for heterogeneity of 20.0 on 19 degrees of freedom, $p > 0.2$. However, this is not so much evidence of homogeneity of relative risks as it is reflective of the high unreliability of the individual relative risks. For 8 of the 20 relative risks shown, the upper limit on the relative risk exceeds the lower limit by a factor of about 10 or more, that factor attaining a value of 57 in one instance.

Blot and Fraumeni express concern about other long term consequences of passive smoking, particularly in connection with coronary artery disease. They cite a report by Garland et al. [8] who initially reported a relative risk due to passive smoking of death from ischemic heart disease of 14.9, but seem unaware that the estimate of 14.9 has been revised downward to 2.7. In the report of the National Research Council [9], which I will be discussing below, there is awareness of the downward revision, but not of the fact that the suggestive significance of $p < 0.10$ is lost and becomes $p < 0.20$.

That lung cancer may aggregate in families is also of concern to Blot and Fraumeni, who cite Ooi et al. on the subject [10]. Elsewhere, and yet to appear, I have suggested that apparent familial aggregation, in the instance breast cancer, may be a reflection of an awareness bias rather than of true familial aggregation [11]. If information about relatives is not collected more directly, the apparent aggregation based on reports from the Index case may only reflect heightened knowledge by such cases of similar illnesses about relatives. But the report by Ooi et al. is another instance, like that of Garland et al., in which there has been unreliable statistical evaluation. Thus, Ooi et al. initially reported that the lung cancer risk increased eighteen-fold per 10-year age increase. By letter in the October 1986 issue of the Journal of the National Cancer Institute they have revised that factor downwards, giving separate factors for each 10-year age interval. From age 50 to age 60, the factor is now reported at only 2.9.

2508012365

The Report of the Committee on Passive Smoking, Board on Environmental Studies and Toxicology, National Research Council [9]

I have chosen to discuss the epidemiologic aspects of this Report separately, since it is essentially the definitive work on current knowledge on environmental tobacco smoke. A member of the committee was Nicholas Wald, senior author of one of the articles discussed above. The report contains a technical appendix which largely duplicates the appendix in the article by Wald et al. and also repeats, with minor variations, the data of Wald et al. The body of the report itself contains those same data, but recast differently, and it is the same 13 studies, with 20 relative risk values, which underlie the epidemiologic aspects of the Committee Report.

There are a great variety of issues which the Committee Report goes into, whether physiochemistry, toxicology, assessment of exposures, use of questionnaires, exposure-dose relationships, etc. But my concern at this time is the epidemiology. There could be a point to estimating the annual number of lung cancer deaths in the United States due to passive smoking, but that would have to be on the presumption that passive smoking does play a causative role.

However, the Committee Report is quite restrained in its findings and leaves open the question of whether anything has been established. If the apparent relative risk is significantly greater than unity, the excess cannot be fully explained away by certain biases considered. However, whether there is statistical significance in view of those biases is not addressed.

From dosimetric considerations, the Report suggests that the excess risk of lung cancer due to environmental tobacco smoke should be 1% of the excess risk due to active smoking. This leads to a relative risk of 1.14 for men, perhaps less for women. From the epidemiologic data, the summary relative risk is 1.34, but it is brought out that for United States studies only the relative risk would be only 1.14. If only large studies are considered, the overall relative risk would be 1.32.

Next addressed by the Report is the effect of biases, particularly the bias associated with the false reporting of individuals that they were not (or never have been) smokers. This leads to a lowering of the estimated relative risk of 1.34 (or 1.30 to 1.34) to 1.15. But note that on this same basis, Wald et al. were willing to reduce an apparent relative risk of 1.35 only slightly, to 1.30.

Yet another adjustment is made. If non-smokers are not exposed to environmental tobacco smoke at home, they might still be exposed to it away from home. An upward adjustment of 8% on account of this yields $1.15 \times 1.08 = 1.24$. This contrasts with the upward adjustment of 18% made by Wald et al., who calculated $1.30 \times 1.18 = 1.53$. The Committee Report differs markedly from the separate report made by one of its own members.

In discussing Wald et al. I suggested that the upward modification they have presupposed a positive role for passive smoking. This same thing is true for the 8% upward adjustment in the Committee Report. For purposes of evaluating the statistical significance of the findings, the relative risk should be taken as 1.15, though the value of 1.24 might be appropriate for assessing the toll in excess lung cancer due to passive smoking assuming that there is causality. With the United States studies indicating an unadjusted relative risk of only 1.14 rather than 1.34, both the 1.15 and the 1.24 might be sharply lowered if intended to apply only to the United States.

But let me stay with the relative risk of 1.15 prior to the 8% upward adjustment. Is that relative risk significantly in excess of 1.00? I suspect not. And even the question of bias remains open. Both in the Committee Report and in the article by Wald et al., the only

2503012366

biases factored in were just those that would fit into neat mathematical formulas. More subtle biases or ones that had not been thought of did not get in. I gave an example above of the use by Garfinkel et al. of the responses by sons and daughters of the level of smoking by the fathers.

I might even speculate about publishing bias. If an investigator got a weakly or insignificantly negative result for the role of passive smoking in lung cancer, would he bother submitting it for publication? And if he did, would it be accepted for publication? Postulating this kind of bias is not necessary for establishing that the 1.15 relative risk is likely not significant. But I bring it up in connection with a tendency I see towards accepting uncritically or less critically manuscripts which are on the right side of the fence on the issue of passive smoking. A particular example was the publication of the article by Garland et al. on passive smoking and ischemic heart disease mortality, the claims of which fell apart on scrutiny.

Let me bring up now another thought. Some time ago the possibility of subtle or not-so-subtle biases in case-control or other epidemiologic investigations was so much a matter of concern that it was suggested that unless the relative risk were at least 2.0, any increase in risk should not be accepted. Perhaps we can do better now and might employ a less restrictive criterion.

But I can see no relaxation to the point of accepting the relative risks now observed for passive smoking in lung cancer. What we must accept is that it is unlikely that any epidemiologic investigation has been on can be mounted which would establish a causal role for passive smoking in lung cancer. Those who believe such a role exists should continue to believe as much, and might even hazard estimates as to the resulting toll in deaths and disease, with others allowed to hold contrary beliefs. What would be incorrect would be to claim that epidemiologic studies have established the correctness of the belief.

If epidemiologic investigations cannot establish a role for passive smoking, the best we can do is to make suppositions estimates of how great that role may be - and such suppositions estimates can be too high if any of the underlying supposals are false. One supposal would be that the dosage response curve is linear through the origin, another that some particular biochemical measure, say level of cotinine, is a proper measure of the equivalent exposure to cigarettes of passive smoking. And, I point out, there could be the assumption that the temperature at which tobacco smoke is inhaled is not relevant, though I would think that fresh hot smoke would be more active than stale smoke.

With this thought in mind, we can pick up some clues from the report of Jarvis et al. who, after excluding "deceivers", report average cotinine levels in plasma, saliva, and urine of 100 non-smokers to be at 0.55%, 0.55% and 0.364 respectively of those levels from 94 smokers. Let us take it at 0.5%. If the average cigarette smoker has a relative risk for lung cancer of 10.0 (enhancement of 900%, though the enhancement may be 1,400% for very active smokers), this would put the enhanced risk due to environmental tobacco smoke at 4.5%, for a relative risk of 1.045 (it would be 1.07 using the 1,400% enhancement for very active smokers). That relative risk, 1.045, would encompass both passive smoking at home and away from home, including individuals not exposed to passive smoking at home.

What matters, however, relative to the conduct of epidemiologic studies on the subject, is the differential in relative risk between those knowingly exposed to passive smoking and those who believe themselves unexposed. From data available in Jarvis et al., it would appear that those seemingly not exposed to passive smoke (46 in number) nevertheless have a relative risk of about 1.02. For the 54 non-smokers claimed to be actually exposed to passive smoking, the relative risk based on cotinine levels would, in

2503012367

similar manner, be 1.07. Compared then to seemingly non-exposed to passive smoking, the calculated relative risk for the known exposed to passive smoking would be 1.05. That small increase in relative risk just would not show up on any epidemiologic investigation and would be submerged, in any case, by other very likely biases. The National Research Council report had suggested a relative risk, based on dosimetric considerations, of 1.14, but on the assumption that enhancement in risk due to an active smoking was 1,400%. An enhancement of 900% would have led them to anticipated relative risk of 1.09. But whether we use 1.05, 1.09, or 1.14, the effect would still be undetectable.

As a last point, I raise the issue of passive smoking effects on children. If parents can be shamed into not exposing their children to passive smoking, this is all well and good, even if the supporting basis is unsound. I note that the ill effects arise mostly in early childhood, and have two questions. Have the passive smoking effects been isolated from effects due to mother's smoking prior to the child's birth? To what extent has account been taken that cigarette smoking concentrates in families with lower socio-economic status, as evidenced by lower educational level and more unemployment etc. Rona et al. also brought in the factor of overcrowding at home in their report that passive smoking resulted in some small reduction in the stature of children [12]. But even Rona et al. failed to take properly into account, as I have suggested, the role of some of these important factors on smoking rates in their evaluation [13].

What with subtle biases, not so subtle biases, and even extravagant errors, one should not accept too readily claimed demonstrations of ill effects of passive smoking. Passive smoking has been the favorite whipping boy of epidemiologists for too long already. The public is entitled not to be unnecessarily exposed to environmental tobacco smoke but any panic is unjustified.

References

1. Blot JB, Fraumeni JF Jr (1986) Guest editorial. Passive smoking and lung cancer. *Journal of the National Cancer Institute* 77:993-1,000
2. Wald NJ, Nanchahal K, Thompson SG, Cuckle HS (1986) Contemporary theme. Does breathing other people's tobacco smoke cause lung cancer? *Br Med J* 293:1,217-22
3. Hirayama T (1981) Non-smoking wives of heavy smokers have a higher risk of lung cancer: A study from Japan. *Br Med J* 282:183-5
4. Mantel N (1983) Guest editorial. Epidemiologic investigations: Care in conduct, care in analysis and care in reporting. *J Cancer Res Clin Oncol* 105:113-6
5. Jarvis M, Tunstall-Pedoe H, Feyerabend C, Vesey C, Saloojee Y (1984) Biochemical markers of smoke absorption and self reported exposure to passive smoking. *Journal of Epidemiology and Community Health* 38:335-9
6. Garfinkel L, Auerbach O, Joubert L (1985) Involuntary smoking and lung cancer: A case-control study. *Journal of the National Cancer Institute* 75:463-9
7. Mantel N (1986) Letter. Involuntary smoking and lung cancer - some object lessons. *Journal of the National Cancer Institute* 76:1,261-3
8. Garland C, Barrett-Connor E, Suarez L, Criqui M, Wingard D (1985) Effects of passive smoking on ischemic heart disease mortality of non-smokers: A prospective study. *Am J Epidemiol* 121:645-50
9. National Research Council (1986) Committee on passive smoking, board on environmental studies and toxicology. *Environmental tobacco smoke - measuring exposures and assessing health effects*. National Academy Press, Washington, DC
10. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H (1986) Increased familial risk for lung cancer. *Journal of the National Cancer Institute* 76:217-22

2503012368

What Is the Epidemiologic Evidence for a Passive Smoking-Lung Cancer Association? 347

11. Mantel N (1989) Letter: Familial breast cancer and the awareness bias. Am J Epidemiol (in press)
12. Rona RJ, Chinn S, Florey C (1985) Exposure to cigarette smoking and children's growth. Int J Epidemiol 14:402-9
13. Mantel N (1986) Letter: Does passive smoking stunt the growth of children? Int J Epidemiol 15:427-8

2503012369

4

2503012370

GLOSSARY

Acute: Having a short course; of short duration.

Animal study: A controlled laboratory experiment in which animals are exposed to an agent and the biological effects of this exposure are assessed. The exposure may be via food or water (ingestion), by injection, by external application or by inhalation. Typical effects that might be measured are tumor incidence or tissue and organ changes.

Bias: Regarding epidemiologic studies, the operation of factors in a study's design or execution that erroneously lead to the appearance of a stronger or weaker association between the agent in question and disease than in fact exists.

Bioassay: The determination of the activity of a sample of an agent by noting its effect on a live animal or an isolated organ preparation.

Carcinogen: A substance or agent designated as capable of producing or initiating cancer.

Carcinogen classification system: A system for stratifying the weight of evidence for human carcinogenicity, for example, the system followed by the EPA. The EPA system consists of

2503012371

the following levels: Group A -- carcinogenic to humans; Group B -- probably carcinogenic to humans; Group C -- possibly carcinogenic to humans; Group D -- not classifiable as to human carcinogenicity; and Group E -- evidence of non-carcinogenicity for humans.¹

Case-control study: A type of epidemiologic study which compares diseased persons (cases) with nondiseased persons (controls) in association with a common exposure to an agent.

Chronic: Persisting over a long period of time. Regarding animal studies, refers to administration of the test substance over a period of several weeks or months.

Cohort study: An epidemiologic study which examines the development of a disease in a group (cohort) of persons who are currently free of the disease. May assess exposure either prospectively or retrospectively.

Confounding: As applied to epidemiologic studies, the situation in which the relationship between an agent and a disease appears stronger or weaker than it truly is due to the influence of another unknown or unrecognized factor. In

1. The definitions for carcinogen classification system, dose-response assessment, exposure assessment, hazard identification, risk assessment, risk characterization and weight of evidence are taken from the EPA's 1986 "Guidelines for Carcinogen Risk Assessment," 51 Fed. Reg. 185, 33992-34003.

2503012372

confounding, the agent under consideration is associated with another agent (a confounding factor, or confounder) which is itself associated with either an increase or decrease in the incidence of the disease.

Dose-response assessment: Part of a risk assessment. Defines the relationship between the dose of an agent and the probability of induction of a carcinogenic effect.

Environmental tobacco smoke (ETS): Consists of smoke originating from the smoldering end of a tobacco product between puffs, e.g., sidestream smoke, and of smoke exhaled by the smoker. The components are released into the environment where they are diluted by ambient air and undergo changes related to aging over time.

Epidemiology: The branch of science concerned with the patterns of disease in human populations and the various factors that influence these patterns.

Exposure assessment: Part of a risk assessment. Identifies populations exposed to the agent, describes their composition and size, and presents the types, magnitudes, frequencies and durations of exposure to the agent.

2503012373

Hazard identification: Part of a risk assessment. A qualitative assessment of risk, dealing with the process of determining whether exposure to an agent has the potential to increase the incidence of cancer. It qualitatively answers the question of how likely an agent is to be a human carcinogen.

In vitro: Literally, within glass; used to refer to laboratory procedures conducted in a test tube or similar location, often involving preparations of cells or tissues.

In vivo: Literally, within the living body; used to refer to laboratory procedures utilizing live animals.

Mainstream smoke (MS): Tobacco smoke drawn through the butt end of a cigarette.

Meta-analysis: A statistical technique for combining studies into a single analysis, designed to increase the ability to statistically detect an association if such an association is present.

Mutagen: An agent that tends to increase the frequency or extent of mutation, i.e., physical or biochemical changes in the genetic material of an organism.

2503012374

Binder, R., et al., "Importance of the Indoor Environment in Air Pollution Exposure," Arch Environ Health 31(6): 277-279, 1976.

Pharmacokinetics: The study of the action of chemical substances in the body over a period of time, including the processes of absorption, distribution, metabolism and excretion.

Relative risk: The ratio of the incidence rate of a disease among individuals exposed to a particular risk factor to the incidence rate among unexposed individuals.

Risk assessment: The determination of adverse health consequences from exposure to toxic agents. [Will be carried out independently from considerations of the consequences of regulatory action.] Includes one or more of the following components: hazard identification, dose-response assessment, exposure assessment and risk characterization.

Risk characterization: Part of a risk assessment. Combines the results of exposure assessment and dose-response assessment to estimate a carcinogenic risk in quantitative terms.

Risk management: A combination of risk assessment with the directives of regulatory legislation, together with socioeconomic, technical, political and other considerations,

2503012375

to reach a decision as to whether or how much to control future exposure to suspected toxic agents.

Short-term tests: In vitro (performed on cells or tissue cultures) tests for mutations, including tests for chromosome aberrations, DNA damage/repair and other transformations which provide supportive evidence of cellular changes and may give information on carcinogenic mechanisms.

Sidestream smoke (SS): Smoke originating from the smoldering end of a tobacco product between puffs.

Statistical significance: A procedure to quantify the probability that an observed outcome, e.g., an association between an exposure and a disease endpoint, arose from random variation alone. The scientific community often uses 5% as a standard level at which data are accepted as occurring other than by chance. This means that there is a 95% probability that the results are not attributable to chance.

Toxicology: The scientific study of poisons, their actions, their detection and the treatment of the conditions produced by them.

Weight of evidence: A framework utilized by the EPA for judging the likelihood that an agent is a human carcinogen. Three major steps are involved: (1) characterization of evidence

from human studies and from animal studies, individually; (2) combination of the characterizations of these two types of data into an indication of the overall weight of evidence; and (3) evaluation of all supporting information to determine if the overall weight of evidence should be modified. [See also definition for carcinogen classification system.]

10380571

2503012377

DEFINITIONS

In medical research, there are two major types of studies: experimental studies and observational studies.

An Experimental Study requires that the members of a study population be assigned to either a treatment or control group. The treated and untreated groups are then followed prospectively to see whether the two groups subsequently differ in their disease experience.

Def: An Observational Study is one in which the treatment or exposure of interest is not assigned but instead occurs by choice or by happenstance.

The types of observational studies are the case report, the cross-sectional study, the ecologic study, the case-control study, and the cohort study (often called prospective).

A Case Report is strictly speaking not a scientific study but a description of a small number of persons with an unusual disease or an unusual change in their disease status.

A Cross-sectional Study reports the characteristics of a group of people at one point in time or a snapshot of their health picture.

The Ecologic Study uses data that are routinely collected (such as air pollution data) to study the occurrence of disease among groups of people. For example, heart disease incidence may be studied in a group of people where information on national dietary habits are known.

A Case-control Study or Retrospective Study is one that begins with study subjects who have the disease of interest and a comparison group without the disease. The previous exposures of both groups are investigated.

A Cohort Study or Prospective Study is one in which the researcher starts with one group of persons exposed to a factor of interest and another comparable group that is unexposed. These groups are observed at a later time to see whether they have developed differences which might be attributable to their different exposures.

A Confounder is a factor which confuses the correct interpretation of the data relating to a suspect and disease. The confounding factor acts by being associated both with the exposure and the disease in a way that makes the exposure and the disease seem to be related. An example, which was published in 1978, related jet plane noise with an increased death rate. Upon reexamination, it was found that persons exposed to jet noise lived in devalued housing close to airports and were of a less fortunate

2503012379

socioeconomic strata. When a proper analysis of these other factors was performed, jet noise was found to have no association with increased mortality.

Confounding is the process by which noncausal associations between two factors is produced by any association with a third factor known as the confounder.

Bias is nonrandom error. Not related to the word bias used in the sense of prejudice.

Risk (or absolute risk) is expressed as a death rate or disease rate.

Relative Risk is the ratio or quotient of two risks or absolute risks. It is also known as a risk ratio.

Odds Ratio is a measure of risk usually obtained from case-control studies and mathematically close to relative risk.

p-value is a statistical estimate of the probability that a finding is due to chance. By convention, a finding with a p-value less than 5%, or sometimes 1%, is called statistically significant.

2503012380

Statistical Association. Two factors are statistically associated when there is a tendency for the factors to occur together or to change together. The observed relationship is the statistical association, measured in many ways.

2503012381

2503012382

A DICTIONARY OF
EPIDEMIOLOGY
SECOND EDITION

*Edited for the
International Epidemiological Association
by*

John M. Last

2503012383

New York Oxford Toronto
OXFORD UNIVERSITY PRESS
1988

Oxford University Press
Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Petaling Java Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland
and associated companies in
Berlin Ibadan

Copyright © 1988 by International Epidemiological Association, Inc.

Published by Oxford University Press, Inc.,
200 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

A Dictionary of epidemiology.
Includes bibliographies.

I. Epidemiology—Dictionaries. I. Last, John M., 1926–
II. International Epidemiological Association.
[DNLM: I. Epidemiology—dictionaries.
WA 13 D559]
RA651.D59 1988 614'.03'21 87-91409
ISBN 0-19-505480-6
ISBN 0-19-505481-4 (pbk.)

2503012384

24681097531

Printed in the United States of America
on acid-free paper

Foreword

The International Epidemiological Association is extremely pleased that the *Dictionary of Epidemiology* has been so successful that a second edition has been demanded. As one of the Association's aims is to "spread the message," this work is an example of "what we call it." Only if we all understand the same thing when a particular term is used will the aim of the Association be capable of being fulfilled. This dictionary is fundamental to this objective.

W. W. Holland, MD FRCGP FRCP FFCM
President, International Epidemiological Association

Preface

This dictionary, appearing now in its second edition, is an attempt to bring some order to the occasionally chaotic nomenclature of epidemiology. It is intended for all who are interested in epidemiology, especially those who are beginning to study the subject, those whose first language is not English, and those from other fields who need to know the terms epidemiologists use.

Like all rapidly expanding sciences, epidemiology has been confounded by the proliferation of words and phrases to describe its concepts, principles, methods, and procedures. The creation of new terms and disagreement about the meaning of old ones can confuse beginners and established epidemiologists alike.

Remarks by users of the first edition have reinforced the view that the boundaries should be wide rather than narrow, that the language should be simple, that some terms many epidemiologists think everyone already knows should be included. The second edition is larger than the first, partly for this reason, and because terms omitted from the first edition have been included and many old entries expanded.

The dictionary is not an index of permitted and proscribed usage. I hope that it is authoritative without being authoritarian. Where synonyms exist, the definition appears under the most commonly used of these, but preference for one term over another is not necessarily implied. In a few instances, the use of a term is deprecated. Some terms that are properly described as slang or jargon have been included because they are widely used and their meaning is not always clear from the context. Murphy's description of jargon is worth recalling: "obscure and/or pretentious language, circumlocutions, invented meanings, and pomposity delighted in for its own sake."

There was disagreement among the contributors to this edition about including certain acronyms and eponyms. An acronym is a word made up of letters from two or more other words, e.g. ANOVA for analysis of variance, or from initial letters, e.g.

2503012385

WHO for World Health Organization. All lay and technical vocabularies contain acronyms; epidemiology has its fair share. By convention, acronyms are spelt out the first time they appear in a text, and, if they are numerous, considerate editors sometimes supply a glossary, or at least list the acronyms along with the words for which they stand in an index. Although this dictionary is not the place for extensive mention of acronyms, a few appeared in the first edition, and a few more appear here.

Eponyms, the attachment of personal or place names to concepts, diseases, methods or specific studies, also occur often enough in published papers and books for us to recognize that beginners need some guidance to the meaning of those most widely used. Some appeared in the first edition, and a few have been added to the second—though again this dictionary is not the proper place for a full glossary of epidemiological eponyms (where would such a glossary end!).

As was the case with the first edition, a large number of epidemiologists from many countries have participated in this revision. The original modest notices in a couple of journals and a few casual remarks among friends produced a mailing list of some forty persons, mainly in North America and the United Kingdom. The mailing list rapidly grew until, by the fifth round of correspondence in December 1986, there were 108 correspondents in 25 countries. The list continued to grow after this fifth and final round; but the published roster of names that follows this preface is both more and less than the number of active participants. Some seemingly inquired just from curiosity and played no further part. Others wrote lengthy and often vigorously argumentative comments and suggestions expressing not only their own views but those of colleagues in their academic department or institution—in one instance, colleagues elsewhere in that nation.

In addition to extensive comments from these correspondents, I have made good use of other technical dictionaries and glossaries in compiling this revision. All of these are listed in the bibliography, and many are also to be found in footnotes that follow specific entries.

The compilers of dictionaries must exercise the greatest care in the choice of words and in their arrangement. Most entries in this dictionary have been repeatedly discussed with many contributors, and in nearly all instances the wording has been agreed upon by all; on the rare occasions when agreement eluded us, the final decision was mine alone. Therefore, I accept full responsibility for the deficiencies in the finished product.

The work has been sponsored by the International Epi-

miological Association, which provided partial travel support for me to attend two meetings in 1986; further support was provided by the McLean Foundation and the Milbank Memorial Fund. All royalties from the sale of this edition, like those from the first edition, will go to the International Epidemiological Association.

Finally, I thank Jeffrey House of Oxford University Press for helpful advice and encouragement.

Ottawa, Canada
November 1987

J. M. L.

Contributing Editors

J. H. ABRAMSON
Jerusalem, Israel

URSULA ACKERMAN-LIEBRICH
Basel, Switzerland

ROBERT ALLARD
Montreal, Quebec, Canada

JOHN C. BAILAR III
Washington, DC, USA

CHRISTOPHER BALDOCK
Brisbane, Queensland, Australia

ROBERTO G. BARUZZI
São Paulo, Brazil

ABRAM S. BENENSON
San Diego, CA, USA

ROGER BERNARD
Geneva, Switzerland

JEAN-FRANÇOIS BOIVIN
Montreal, Quebec, Canada

BERNARD J. BRABIN
Madang, Papua New Guinea

C. RALPH BUNCHER
Cincinnati, OH, USA

BEVERLEY CARLSON
New York, NY, USA

JAMES CHIN
Berkeley, CA, USA

MICHEL COLEMAN
Oxford, England

L. CAYOLLA DA MOTTA
Lisbon, Portugal

GARETH DAVIES
New Haw, Surrey, England

RICHARD DICKER
Atlanta, GA, USA

ALVAN R. FEINSTEIN
New Haven, CT, USA

DAVID FINNEY
Edinburgh, Scotland

JOSEPH L. FLEISS
New York, NY, USA

GARY D. FRIEDMAN
Oakland, CA, USA

MICHAEL GARRAWAY
Edinburgh, Scotland

SANDER GREENLAND
Los Angeles, CA, USA

TEE GUIDOTTI
Edmonton, Alberta, Canada

WALTER W. HOLLAND
London, England

BARBARA HULKA
Chapel Hill, NC, USA

MICHEL İBRAHİM
Chapel Hill, NC, USA

LESLIE M. IRWIG
Sydney, NSW, Australia

MILOS JENICEK
Montreal, Quebec, Canada

L. KARHAUSEN
Luxembourg, Luxembourg

2503012387

contributing editors

HENK LAMBERTS
Amsterdam, the Netherlands

JOHN M. LAST
Ottawa, Ontario, Canada

DALE LAWRENCE
Atlanta, GA, USA

DAVID E. LILIENFELD
New York, NY, USA

GENEVIEÈE LOSLIER
Hudson, Quebec, Canada

ROBERT MACLENNAN
Brisbane, Queensland, Australia

MARGARET F. McCANN
Chapel Hill, NC, USA

ANTHONY B. MILLER
Toronto, Ontario, Canada

KIUMARSS NASSERI
Teheran, Iran

JOHN S. NEUBERGER
Kansas City, KA, USA

NORMAN D. NOAH
London, England

ROBERT OSEASOON
San Antonio, TX, USA

HARRIS PASTIDES
Amherst, MA, USA

MIGUEL PORTA
Barcelona, Spain

DAVID ROBINSON
Geneva, Switzerland

GEOFFREY A. ROSE
London, England

xii

KENNETH J. ROTHRMAN
Boston, MA, USA

JAMES J. SCHLESSELMAN
Bethesda, MD, USA

B.A. SOUTHGATE
London, England

CLAUDE STROHMENGER
Ottawa, Ontario, Canada

IAN SUTHERLAND
Cambridge, England

MERVYN SUSNER
New York, NY, USA

A.V. SWAN
London, England

RODOLFO SARACCI
Lyon, France

MICHEL THURIAUX
Copenhagen, Denmark

B. TOMA
Maisons-Alfort, France

CARL TYLER
Atlanta, GA, USA

ROBERT B. WALLACE
Iowa City, IA, USA

STEPHEN D. WALTER
Hamilton, Ontario, Canada

KERR L. WHITE
Stanardsville, VA, USA

DONALD WICLE
Ottawa, Ontario, Canada

CORRESPONDING EDITORS

ERIK ALLANDER
Huddinge, Sweden

ALBERTA ALZATE
Cali, Colombia

GEOFFREY A. ANDERSON
Ottawa, Ontario, Canada

MARY JANE ASHLEY
Toronto, Ontario, Canada

xiii

corresponding editors

R.S. BHOPAL
Glasgow, Scotland

PATRICIA A. BUFFLER
Houston, TX, USA

ARVIND A. CARPENTER
Oak Ridge, TN, USA

CARL J. CASPERSON
Atlanta, GA, USA

DENIS CHARPIN
Marseilles, France

GERALD R. CHASE
Denver, CO, USA

EMIL E. CRISTOFANO
Akron, OH, USA

ORESTES FAGET CEPERO
Havana, Cuba

ANNE HERSEY COULSON
Los Angeles, CA, USA

DAVID CUNDIFF
Trenton, NJ, USA

ROGER DETELS
Los Angeles, CA, USA

RODGER DOYLE
Buffalo, NY, USA

GERARD DUBOIS
Paris, France

JACQUELINE FABIA
Quebec City, Quebec, Canada

G.I. FORBES
Edinburgh, Scotland

LINA FORCIER
Sydney, NSW, Australia

EDUARDO FRANCO
Sao Paulo, Brazil

JACK FROOM
Stony Brook, NY, USA

TRULS GEDDE-DAHL
Oslo, Norway

O.N. GILL
London, England

DAVID GOLDSMITH
Edmonton, Alberta, Canada

PATRICIA GRAVES
Madang, Papua New Guinea

VINCENT GUINEE
Houston, TX, USA

MATTI HAKAMA
Tampere, Finland

A. SCOTT HENDERSON
Canberra, ACT, Australia

CATHERINE HILL
Villejuif, France

ERNEST B. HOOK
Albany, NY, USA

J. JANCAR
Bristol, England

FINN KAMPER-JORGENSEN
Copenhagen, Denmark

H.W. KANIS
Lelystad, the Netherlands

JENNIFER KELSEY
New York, NY, USA

MARY CLAIRE KING
Berkeley, CA, USA

MAURICE KING
Leeds, England

GEORGE KNOX
Birmingham, England

JESS KRAUS
Los Angeles, CA, USA

H. OLIVER LANCASTER
Sydney, NSW, Australia

F.D.K. LIDDELL
Montreal, Quebec, Canada

IAN McDOWELL
Ottawa, Ontario, Canada

H. MICHAEL MAETZ
Birmingham, AL, USA

Luís MAGAO
Lisbon, Portugal

2503012388

corresponding editors

J.S. MALHI
Brighton, England

DILEEP V. MAVALANKER
Ahmedabad, India

DAVID MORRIS
London, England

C.S. MUIR
Lyon, France

ENRIQUE NÁJERA
Seville, Spain

KATE O'CONNOR
Ottawa, Ontario, Canada

H. TUNSTALL PEDOE
Dundee, Scotland

I. PLESKO
Bratislava, Czechoslovakia

PAULA RANTIKALLIO
Oulu, Finland

R.A. ROBINSON
St. Paul, MN, USA

ROGER ROCHAT
Atlanta, GA, USA

JEFFREY ROSEMAN
Birmingham, AL, USA

PHILIP ROSS
Honolulu, HI, USA

M. ŠARIĆ
Zabreb, Yugoslavia

BJORN SMEDBY
Uppsala, Sweden

ANDREU SEGURA
Barcelona, Spain

xiv

COLIN SOSKOLNE
Edmonton, Alberta, Canada

PAULA STEWART
Ottawa, Ontario, Canada

ROBERT SPASOFF
Ottawa, Ontario, Canada

D. STRACHILOV
Sofia, Bulgaria

HUGH TILSON
Research Triangle Park, NC, USA

TOSHIO TOYAMA
Tokyo, Japan

EDWARD J. TRAPIDO
Miami, FL, USA

STJEPAN VIDACEK
Zagreb, Yugoslavia

ANNE WALLING
Wichita, KA, USA

ROBERT WEST
Cardiff, Wales

JAN WIENPAHL
Los Angeles, CA, USA

WU XI-KE
Hefei, Anhui, China

M.J. WYSOCKI
Warsaw, Poland

FARAT YUSUF
North Ryde, NSW, Australia

FRED ZERFAS
Los Angeles, CA, USA

SHUXIAN ZU
Hefei, Anhui, China

A Dictionary of Epidemiology

2503012389

A

ABORTION RATE The estimated annual number of abortions per 1000 women of reproductive age (usually defined as age 15-44).

ABORTION RATIO The estimated number of abortions per 100 live births in a given year.
ABSCISSA The distance along the horizontal coordinate or *x* axis, of a point *P* from the vertical or *y* axis of a graph. See also **AXIS, GRAPH, ORDINATE**.

ABSOLUTE RISK Usually this term means the observed or calculated risk of an event in a population under study, as contrasted with the relative risk. Sometimes, however, it is a synonym for attributable fraction, excess risk, or risk difference; because of the inconsistency, this term should be avoided. See also **RISK**.

ACCEPTABLE RISK The risk that has minimal detrimental effects, or for which the benefits outweigh the potential hazards. Epidemiologic study has provided data for calculation of risks associated with many medical procedures and also with occupational and environmental exposures; these data are used, for instance, in **CLINICAL DECISION ANALYSIS**.

ACCURACY The degree to which a measurement, or an estimate based on measurements, represents the true value of the attribute that is being measured. See also **MEASUREMENT, PROBLEMS WITH TERMINOLOGY**.

ACQUAINTANCE NETWORK Group of persons in contact or communication among whom transmission of an infectious agent and of knowledge, attitudes, and values is possible, and whose social interaction may have health implications. See also **TRANSMISSION OF INFECTION**.

ACQUIRED IMMUNODEFICIENCY SYNDROME (Syn: acquired immune deficiency syndrome) (AIDS) For surveillance purposes, the Centers for Disease Control, Atlanta, Georgia,¹ define a case of AIDS as an illness characterized by (1) one or more of a group of opportunistic or indicator diseases that are indicative of underlying cellular immunodeficiency; (2) absence of all known underlying causes of cellular immunodeficiency and absence of all other causes of reduced resistance to opportunistic or indicator diseases. Additional criteria are serum positive for HIV antibody, positive culture for HIV, and reduction of T4 "helper" lymphocytes.

The opportunistic or indicator diseases associated with AIDS include certain protozoal and helminth infections, notably *Pneumocystis carinii* pneumonia and toxoplasmosis; fungal infections, notably candidiasis of esophagus, trachea, bronchi or lungs and cryptococcosis, especially affecting the central nervous system; bacterial infections, notably with certain mycobacteria; viral infections, notably cytomegalovirus and herpes simplex; and cancer, notably Kaposi's sarcoma and lymphoma limited to the brain.

AIDS-related complex (ARC) is the combination of HIV positive test with lymph-

2503012390

adenopathy and persistent low fever but without immunodeficiency or opportunistic diseases.

¹1987 Revision of case definition of AIDS for surveillance purposes. MMWR 36, 1S:4S-9S, 1987.

ACTIVITIES OF DAILY LIVING (ADL) SCALE A scale devised by Katz and others¹ to score physical ability/disability; used to measure outcomes of interventions for various chronic disabling conditions such as arthritis. The scale is based on scores for responses to questions about mobility, self-care, grooming, etc. This was the first widely used scale of this type; others, mostly refinements or variations of the ADL scale, have since been developed.

¹Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: Studies of illness in the aged. The index of ADL, a standardized measure of biological function. JAMA 185:914-919, 1963.

ACTUARIAL RATE See FORCE OF MORTALITY.

ACTUARIAL TABLE See LIFE TABLE.

ACUTE

1. Referring to a health effect, brief; sometimes loosely used to mean severe.
2. Referring to exposure, brief, intense, or short-term; sometimes specifically referring to brief exposure of high intensity. See also CHRONIC.

ADAPTATION A heritable component of the phenotype which confers an advantage in survival and reproductive success. The process by which organisms adapt to environmental conditions.

ADDITIVE MODEL A model in which the combined effect of several factors is the sum of the effects that would be produced by each of the factors in the absence of the others. For example, if factor X adds x% to risk in the absence of Y, and if factor Y adds y% to risk in the absence of X, an additive model states that the two factors together will add (x+y)% to risk. See also INTERACTION; LINEAR MODEL; MATHEMATICAL MODEL; MULTIPLICATIVE MODEL.

ADJUSTMENT A summarizing procedure for a statistical measure in which the effects of differences in composition of the populations being compared have been minimized by statistical methods. Examples are adjustment by regression analysis and by standardization. Adjustment often is performed on rates or relative risks, commonly because of differing age distributions in populations that are being compared. The mathematical procedure commonly used to adjust rates for age differences is direct or indirect standardization.

ADVERSE REACTION, SIDE EFFECT Any undesirable or unwanted consequence of a preventive, diagnostic, or therapeutic procedure.

AETIOLOGY, AETIOLOGIC See ETIOLOGY, ETIOLOGIC.

AGE DEPENDENCY RATIO See DEPENDENCY RATIO.

AGENT (OF DISEASE) A factor, such as a microorganism, chemical substance, or form of radiation, whose presence, excessive presence, or (in deficiency diseases) relative absence is essential for the occurrence of a disease. A disease may have a single agent, a number of independent alternative agents (at least one of which must be present), or a complex of two or more factors whose combined presence is essential for the development of the disease. See also CAUSALITY; NECESSARY AND SUFFICIENT CAUSE.

AGE-PERIOD COHORT ANALYSIS See COHORT ANALYSIS.

AGE-SEX PYRAMID See POPULATION PYRAMID.

AGE-SEX REGISTER List of all clients or patients of a medical practice or service, classified by age (birthdate) and sex; provides denominator for calculating age- and sex-specific rates.

AGE-SPECIFIC FERTILITY RATE The number of births occurring during a specified pe-

riod to women of a specified age group, divided by the number of person-years lived during that period by women of that age group. When an age-specific fertility rate is calculated for a calendar year, the number of births to women of the specified age is usually divided by the midyear population of women of that age.

AGE-SPECIFIC RATE A rate for a specified age group. The numerator and denominator refer to the same age group.

Example:

$$\text{Age-specific death rate (age 25-34)} = \frac{\text{Number of deaths among residents age 25-34 in an area in a year}}{\text{Average (or midyear) population age 25-34 in the area in that year}} \times 100,000$$

The multiplier (usually 100,000 or 1,000,000) is chosen to produce a rate that can be expressed as a convenient number.

AGE STANDARDIZATION A procedure for adjusting rates, e.g. death rates, designed to minimize the effects of differences in age composition when comparing rates for different populations. See also ADJUSTMENT, STANDARDIZATION.

AGGREGATION BIAS (Syn: ecological bias) See ECOLOGICAL FALLACY.

AGING OF THE POPULATION A demographic term, meaning an increase over time in the proportion of older persons in the population. It does not necessarily imply an increase in life expectancy or that "people are living longer than they used to." The principal determinant of aging in the population has been a decline in the birth rate; when fewer children are born than in prior years, the result, in the absence of a rise in the death rate at higher ages, has been an increase in the proportion of older persons in the population. In developed societies, however, mortality change is becoming a factor: little further mortality reduction can occur in the first half of life, so reductions are beginning to occur in the third and fourth quarters of life, leading to a rise in the proportion of older persons from this cause.

AIRBORNE INFECTION A mechanism of transmission of an infectious agent by particles, dust, or DROPLET NUCLEI suspended in the air. See also TRANSMISSION OF INFECTION.

ALGORITHM Any systematic process that consists of an ordered sequence of steps with each step depending on the outcome of the previous one. The term is commonly used to describe a structured process, for instance, relating to computer programming or to health planning. See also DECISION TREE.

ALGORITHMIC, CLINICAL (Syn: clinical protocol) An explicit description of steps to be taken in patient care in specified circumstances. This approach makes use of branching logic and of all pertinent data, both about the patient and from epidemiologic and other sources, to arrive at decisions that yield maximum benefit and minimum risk.

ALLEL Alternative forms of a gene, occupying the same locus on a chromosome.

ALPHA ERROR See ERROR, TYPE I.

ALPHA LEVEL See SIGNIFICANCE LEVEL.

ANALYSIS OF VARIANCE A statistical technique that isolates and assesses the contribution of categorical independent variables to variation in the mean of a continuous dependent variable. The observations are classified according to their categories for each of the independent variables, and the differences between the categories in their mean values on the dependent variable are estimated and tested for statistical significance.

ANALYTIC STUDY A study designed to examine associations, commonly putative or hypothesized causal relationships. An analytic study is usually concerned with identi-

203012391

fying or measuring the effects of risk factors, or is concerned with the health effects of specific exposure(s). Contrast descriptive study, which does not test hypotheses. The common types of analytic study are CROSS-SECTIONAL, COHORT, and CASE-CONTROL. In an analytic study, individuals in the study population may be classified according to absence or presence (or future development) of specific disease and according to "attributes" that may influence disease occurrence. Attributes may include age, race, sex, other disease(s), genetic, biochemical, and physiological characteristics, economic status, occupation, residence, and various aspects of the environment or personal behavior. See also CASE CONTROL STUDY; COHORT STUDY; CROSS-SECTIONAL STUDY; STUDY DESIGN.

ANIMAL MODEL Study in a population of laboratory animals that uses conditions of animals analogous to conditions of humans to model processes comparable to those that occur in human populations. See also EXPERIMENTAL EPIDEMIOLOGY.

ANTAGONISM Opposite of SYNERGISM. The situation in which the combined effect of two or more factors is smaller than the solitary effect of any one of the factors. In BIOASSAY, the term may be used to refer to the situation when a specified response is produced by exposure to either of two factors but not by exposure to both together.

ANTHROPOMETRY The technique that deals with the measurement of the size, weight, and proportions of the human body.

ANTHROPOPHILIC (adj.) Pertaining to an insect's preference for feeding on humans even when nonhuman hosts are available.

ANTIBODY Protein molecule formed by exposure to a "foreign" or extraneous substance, e.g., invading microorganisms responsible for infection, or active immunization. May also be present as a result of passive transfer from mother to infant, via immune globulin, etc. Antibody has the capacity to bind specifically to the foreign substance (antigen) that elicited its production, thus supplying a mechanism for protection against infectious diseases. Antibody is epidemiologically important because its concentration (titer) can be measured in individuals, and, therefore, in populations. See also SEROEPIDEMIOLOGY.

ANTIGEN A substance (protein, polysaccharide, glycolipid, tissue transplant, etc.) that is capable of inducing specific immune response. Introduction of antigen may be by the invasion of infectious organisms, immunization, inhalation, ingestion, etc.

ANTIGENIC DRIFT This term describes the "evolutionary" changes that take place in the molecular structure of DNA/RNA in micro-organisms during their passage from one host to another. It may be due to recombination, deletion or insertion of genes, to point mutations, or to several of these events. This process has been studied in common viruses, notably the influenza virus.¹ It leads to alteration (usually slow and progressive) in the antigenic composition, and thus in the immunologic responses of individuals and populations to exposure to the micro-organisms concerned. See also ANTIGENIC SHIFT.

¹Palese P, Young JF: Variation of Influenza A, B, and C Viruses. *Science* 215:1468-1473, 1982.

ANTIGENIC SHIFT This term describes mutation, i.e., a sudden change in molecular structure of DNA/RNA in micro-organisms, especially viruses, which produces new strains of the micro-organism. Hosts previously exposed to other strains have little or no acquired immunity. Antigenic shift is believed to be the explanation for the occurrence of strains of the influenza A virus associated with large-scale epidemic and pandemic spread. Antigenic shift is responsible for the susceptibility of host populations to a new strain of influenza virus. See also ANTIGENIC DRIFT.

ANTIGENICITY (Syn: immunogenicity) The ability of agent(s) to produce a systemic or a local immunologic reaction in the host.

ARBOVIRUS A group of taxonomically diverse animal viruses that are unified by an epidemiologic concept, i.e., transmission between vertebrate host organisms by blood-feeding (hematophagous) arthropod vectors such as mosquitoes, ticks, sand flies, and midges. The term is a contraction of *arthropod-borne virus*.

The interaction of arbovirus, vertebrate host(s), and arthropod vector gives this class of infections several unique epidemiologic features. See VECTOR-BORNE INFECTION for definition of terms used to describe these features.

AREA SAMPLING A method of sampling that can be used when the numbers in the population are unknown. The total area to be sampled is divided into subareas, e.g., by means of a grid that produces squares on a map; these subareas are then numbered and sampled, using a table of random numbers. Depending upon circumstances, the population in the sampled areas may first be enumerated, then a second stage of sampling may be conducted.

ARITHMETIC MEAN The sum of all the values in a set of measurements, divided by the number of values in the set.

ARTIFICIAL INTELLIGENCE A branch of computer science in which attempts are made to duplicate human intellectual functions. One application is in diagnosis, in which computer programs are often based upon epidemiologic analyses of data in hospital charts or other clinical records.

ASCERTAINMENT The process of determining what is happening in a population or study group, e.g., family and household composition, occurrence of cases of specific diseases; the latter is also known as case-finding.

ASCERTAINMENT BIAS Systematic failure to represent equally all classes of cases or persons supposed to be represented in a sample. This bias may arise because of the nature of the sources from which persons come, e.g., a specialized clinic; from a diagnostic process influenced by culture, custom, or idiosyncracy; or, for example, in genetic studies, from the statistical chance of selecting from large or small families.

ASSAY The quantitative or qualitative evaluation of a hazardous substance; the results of such an evaluation.

ASSOCIATION (Syn: correlation, [statistical] dependence, relationship) Statistical dependence between two or more events, characteristics, or other variables. An association is present if the probability of occurrence of an event or characteristic, or the quantity of a variable, depends upon the occurrence of one or more other events, the presence of one or more other characteristics, or the quantity of one or more other variables. The association between two variables is described as positive when the occurrence of higher values of a variable is associated with the occurrence of higher values of another variable. In a negative association, the occurrence of higher values of one variable is associated with lower values of the other variable. An association may be fortuitous or may be produced by various other circumstances; the presence of an association does not necessarily imply a causal relationship. If the use of the term "association" is confined to situations in which the relationship between two variables is statistically significant, the terms "statistical association" and "statistically significant association" become tautological. However, ordinary usage is seldom so precise as this. The terms "association" and "relationship" are often used interchangeably.

Associations can be broadly grouped under two headings, symmetrical or non-causal (see below) and asymmetrical or causal.

ASSOCIATION, ASYMMETRICAL (Syn: asymmetrical relationship) The definitive conditions of asymmetrical associations are direction and time. Independent variable X must cause changes in dependent variable Y, and the "causal" variable must precede it.

250301238

"effect." Bradford Hill¹ and others^{2,3} have pointed out that the (subjective) likelihood of a causal relationship is increased by the presence of the following attributes. However, temporality is the only indispensable condition among these.

1. Consistency—The association is consistent if the results are replicated when studied in different settings and by different methods.
2. Strength—This is an expression of the disparity between the frequency with which a factor is found in the disease and the frequency with which it occurs in the absence of the disease. Not to be confused with statistical significance.
3. Specificity—This is established with the limitation of the association to a single putative cause and single effect.
4. Dose-response relationship—This is established when an increased risk or severity in disease occurs with an increased quantity ("dose") or duration of exposure to a factor.
5. Temporality—The exposure to a putative cause always precedes, never follows, the outcome.
6. Biological plausibility—It is desirable that the association agree with current understanding of the response of cells, tissues, organs, and systems to stimuli. This criterion should not be applied rigidly. The association may be new to science or medicine. As Sherlock Holmes advised Dr. Watson, "When you have eliminated the impossible, whatever remains, however improbable, must be the truth."
7. Coherence—The associations should not conflict with the generally known facts of the natural history and biology of disease.
8. Experiment—It is sometimes possible to appeal to experimental, or quasi-experimental evidence, e.g., an observed association leads to some preventive action. Does this action in fact prevent?

See also CAUSALITY; EVANS'S POSTULATES; KOCH'S POSTULATES.

¹Bradford Hill A: The environment and disease: Association or causation. *Proc Roy Soc Med* 58:295-300, 1965.

²Susser MW: Judgment and causal inference. *Am J Epidemiol* 105:1-15, 1977.

³Rothman KJ (Ed): *Causal Inference*. Chestnut Hill, MA: Epidemiology Resources Inc., 1988.

ASSOCIATION, DIRECT Directly associated, i.e., not via a known third variable: $A \rightarrow B$. Refers only to causality.

ASSOCIATION, INDIRECT CAUSAL Two types are distinguished:

1. Association of a factor C with disease A only because both are related to a common underlying factor B.



Alteration of factor C will not produce an alteration in the frequency of disease A unless an alteration in C affects B. It has been suggested that to avoid confusion with the alternative meaning of indirect association, this type should be called "secondary association."

2. Association of a factor C with disease A by means of an intermediate or intervening factor B.



Alteration of factor C would produce an alteration in the frequency of disease A. To avoid confusion, this type should be called "indirect causal association."

ASSOCIATION, SPURIOUS A term, preferably avoided, used with different meanings by different authors. It may refer to artifactual, fortuitous, false secondary, or to all kinds of noncausal associations due to chance, bias, failure to control for extraneous variables, etc.

ASSOCIATION, SYMMETRICAL An association is noncausal if it is symmetrical, as in the statement $F = MA$ (force equals mass times acceleration). This is a noncausal, non-directional expression of the mathematical relationship between the physical properties of force, mass, and velocity. If one side of the equation is changed, then the other must also change to maintain equilibrium.

Although epidemiologists are usually most interested in asymmetrical statements that have direction, the symmetrical equation can be useful. For instance, prevalence can be expressed in terms of incidence and duration in the simple equation, $P = I \times D$. If two of these three elements are known, the third can be derived. See also SYMMETRICAL RELATIONSHIP.

ASSORTATIVE MATING Selection of a mate with preference (or aversion) for a particular genotype, i.e., nonrandom mating.

ASYMMETRICAL ASSOCIATION See ASSOCIATION, ASYMMETRICAL.

ASYMPTOTIC Pertaining to a limiting value, for example, of a dependent variable, when the independent variable approaches zero or infinity. See LARGE SAMPLE METHOD.

ASYMPTOTIC METHOD See LARGE SAMPLE METHOD.

ATTACK RATE Attack rate, or case rate, is a CUMULATIVE INCIDENCE RATE often used for particular groups, observed for limited periods and under special circumstances, as in an epidemic.

The secondary attack rate is the number of cases among contacts occurring within the accepted incubation period following exposure to a primary case, in relation to the total of exposed contacts; the denominator may be restricted to susceptible contacts when determinable.

Infection rate is the incidence of manifest plus inapparent infections, which can be identified, e.g., by SEROEPIDEMIOLOGY.

ATTRIBUTABLE FRACTION (AF) (Syn: attributable proportion) A term sometimes used to refer to the attributable fraction in the population, and sometimes to the attributable fraction among the exposed. See also ATTRIBUTABLE FRACTION (EXPOSED); ATTRIBUTABLE FRACTION (POPULATION).

ATTRIBUTABLE FRACTION (EXPOSED) (Syn: attributable proportion [exposed], attributable risk, etiologic fraction [exposed]). With a given outcome, exposure factor and population, the attributable fraction among the exposed is the proportion by which the incidence rate of the outcome among those exposed would be reduced if the exposure were eliminated. It may be estimated by the formula

$$AF_e = \frac{I_e - I_u}{I_e}$$

where I_e is the incidence rate among the exposed, I_u is the incidence rate among the unexposed; or by the formula

$$AF_e = \frac{RR - 1}{RR}$$

where RR is the rate ratio, I_e/I_u . It is assumed that causes other than the one under investigation have had equal effects on the exposed and unexposed groups.

ATTRIBUTABLE FRACTION (POPULATION) (Syn: attributable proportion [population], etiologic fraction [population], attributable risk). With a given outcome, exposure factor, and population, the attributable fraction among the population is the propor-

tion by which the incidence rate of the outcome in the entire population would be reduced if exposure were eliminated. It may be estimated by the formula

$$AF_p = \frac{I_p - I_u}{I_p}$$

where I_p is the incidence rate in the total population and I_u is the incidence rate among the unexposed; or by the formula

$$\frac{P_e(RR - 1)}{1 + P_e(RR - 1)}$$

where RR is the rate ratio, I_u/I_p . It is assumed that causes other than the one under investigation have had equal effects on the exposed and unexposed groups.

ATTRIBUTABLE NUMBER The number of new occurrences of a specific outcome attributable to an exposure; it may be estimated using the formula

$$AN = \frac{I_e - I_u}{N_e}$$

where I_e is the incidence rate among the exposed, I_u is the incidence rate among the unexposed, and N_e is the number of persons in the exposed population. It is assumed that causes other than the one under investigation have had equal effects on the exposed and unexposed groups.

ATTRIBUTABLE RISK The rate of a disease or other outcome in exposed individuals that can be attributed to the exposure. This measure is derived by subtracting the rate of the outcome (usually incidence or mortality) among the unexposed from the rate among the exposed individuals; it is assumed that causes other than the one under investigation have had equal effects on the exposed and unexposed groups. Unfortunately, this term has been used to denote a number of different concepts, including the attributable fraction in the population, the attributable fraction among the exposed, the population excess rate, and the rate difference. Therefore, it should be defined carefully by all who use it. See also ATTRIBUTABLE FRACTION (EXPOSED); POPULATION EXCESS RATE; ATTRIBUTABLE FRACTION (POPULATION); POPULATION ATTRIBUTABLE RISK; RATE DIFFERENCE.

ATTRIBUTABLE RISK (EXPOSED) This term has been used with different connotations to denote the attributable fraction among the exposed and the excess risk among the exposed. See also ATTRIBUTABLE FRACTION (EXPOSED); RATE DIFFERENCE.

ATTRIBUTABLE RISK (POPULATION) This term has been used with different connotations to denote the attributable fraction in the population and the population excess risk. See also ATTRIBUTABLE FRACTION (POPULATION); POPULATION EXCESS RATE.

ATTRIBUTABLE RISK PERCENT Attributable fraction expressed as a percentage rather than as a proportion.

ATTRIBUTABLE RISK PERCENT (EXPOSED) This is the attributable fraction among the exposed, expressed as a percentage. See also ATTRIBUTABLE FRACTION (EXPOSED).

ATTRIBUTABLE RISK PERCENT (POPULATION) This is the attributable fraction in the population, expressed as a percentage. See also ATTRIBUTABLE FRACTION (POPULATION).

ATTRIBUTE A qualitative characteristic of an individual or item.

AUDIT An examination or review that establishes the extent to which a condition, process, or performance conforms to predetermined standards or criteria.

AUTOPSY DATA Data derived from autopsied deaths, e.g., for study of natural history of disease and trends in frequency of disease. Autopsies are done on nonrandomly selected persons in the population and findings should therefore be generalized only with great caution.

AVERAGE Kendall and Buckland's *Dictionary of Statistical Terms* (4th Edition, 1982) has this to say: "A familiar but elusive concept. Generally an 'average' value purports to represent or to summarize the relevant features of a set of values; and in this sense the term would include the median and the mode. In a more limited sense an 'average' compounds all the values of the set, e.g., in the case of the arithmetic or geometric means. In ordinary usage, 'the average' is often understood to refer to the arithmetic mean." See also MEASURES OF CENTRAL TENDENCY.

AVERAGE LIFE EXPECTANCY See EXPECTATION OF LIFE.

AXIS

1. One of the dimensions of a graph. A two-dimensional graph has two axes, the horizontal or x axis, and the vertical or y axis. Mathematically, there may be more than two axes, and graphs are sometimes drawn with a third dimension; the eye cannot comprehend more than three dimensions.
2. In NOSOLOGY, an axis of classification is the conceptual framework, e.g., etiologic, topographic, psychologic, sociologic. The International Classification of Disease, for example, is multiaxial; the primary axis is topographic (i.e., body systems); secondary axes relate to etiology, manifestations of disease, detail of sites affected, severity, etc.

2503012394

$$P(D|S) = \frac{P(S|D)P(D)}{P(S|D)P(D) + P(S|\bar{D})P(\bar{D})}$$

where D = disease, S = symptom, and \bar{D} = no disease. The formula emphasizes what clinical intuition often overlooks, namely, that the probability of disease given this symptom depends not only on how characteristic that symptom is of the disease but also on how frequent the disease is among the population being served. "If you hear hoof beats in the street, do not look for zebra."

The theorem can also be used for estimating exposure-specific rates from case control studies if there is added information about the overall rate of disease in that population.

Some of the terms in the theorem have special names. The probability of disease given the symptom is called the "posterior probability." It is an estimate of the probability of disease posterior to knowing whether or not the symptom was present. The overall probability of disease among the population or our guess of the probability of disease before knowing of the presence or absence of the symptom is called the "prior probability." The theorem is sometimes presented in terms of the odds of disease before knowing the symptom (prior odds) and after knowing the symptom (posterior odds).

BEHAVIORAL EPIDEMIC An epidemic originating in behavioral patterns (as opposed to invading microorganisms or physical agents). Examples include the dancing manias of the Middle Ages, episodes of mass fainting or convulsions ("hysterical epidemics"), crowd panic, or waves of fashion or enthusiasm. The communicable nature of the behavior is dependent not only on person-to-person transmission of the behavioral pattern but also on group reinforcement (as with smoking, alcohol, or drug use). Behavioral epidemics may be difficult to differentiate from, or may complicate, outbreaks of organic disease, for example, due to contamination of the environment by a toxic substance.

BEHAVIORAL RISK FACTOR A characteristic or behavior that is associated with increased probability of a specified outcome; the term does not imply a causal relationship.

BENCHMARK A slang or jargon term, usually meaning a measurement taken at the outset of a series of measurements of the same variable, sometimes meaning the best or most desirable value of the variable. Because of uncertainty about meaning, the term should not be used.

BENEFIT-COST RATIO The ratio of net present value of measurable benefits to costs. Calculation of a benefit-cost ratio is used to determine the economic feasibility or success of a program.

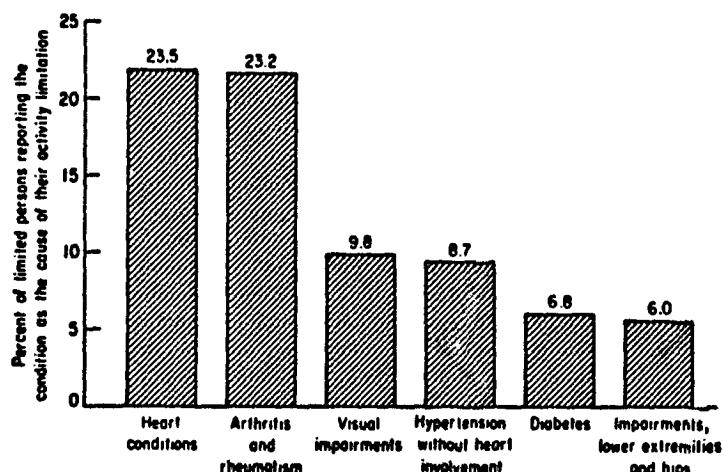
BERNOULLI DISTRIBUTION The probability distribution associated with two mutually exclusive and exhaustive outcomes, e.g., death or survival; a Bernoulli variable is one that has only two possible values, e.g., death or survival. See also BINOMIAL DISTRIBUTION.

BERKSON'S BIAS See BIAS, SELECTION.

BETA ERROR See ERROR, TYPE II.

BIAS Deviation of results or inferences from the truth, or processes leading to such deviation. Any trend in the collection, analysis, interpretation, publication, or review of data that can lead to conclusions that are systematically different from the truth. Among the ways in which deviation from the truth can occur, are the following:

1. Systematic (one-sided) variation of measurements from the true values (syn: systematic error).



Bar diagram. From Susser, Watson, Hopper, 1985.

BAYES' THEOREM A theorem in probability theory named for Thomas Bayes (1702–1761), an English clergyman and mathematician; his *Essay Towards Solving a Problem in the Doctrine of Chances* (1763, published posthumously), contained this theorem. In epidemiology, it is used to obtain the probability of disease in a group of people with some characteristic on the basis of the overall rate of that disease (the prior probability of disease) and of the likelihoods of that characteristic in healthy and diseased individuals. The most familiar application is in CLINICAL DECISION ANALYSIS where it is used for estimating the probability of a particular diagnosis given the appearance of some symptoms or test result. A simplified version of the theorem is

2. Variation of statistical summary measures (means, rates, measures of association, etc.) from their true values as a result of systematic variation of measurements, other flaws in data collection, or flaws in study design or analysis.
3. Deviation of inferences from the truth as a result of flaws in study design, data collection, or the analysis or interpretation of results.
4. A tendency of procedures (in study design, data collection, analysis, interpretation, review or publication) to yield results or conclusions that depart from the truth.
5. Prejudice leading to the conscious or unconscious selection of study procedures that depart from the truth in a particular direction, or to one-sidedness in the interpretation of results.

The term "bias" does not necessarily carry an imputation of prejudice or other subjective factor, such as the experimenter's desire for a particular outcome. This differs from conventional usage in which bias refers to a partisan point of view.

Many varieties of bias have been described.¹

¹Sackett DL: Bias in analytic research. *J Chron Dis* 32:51-63, 1979.

BIAS, ASCERTAINMENT Systematic error, arising from the kind of individuals or patients (e.g., slightly ill, moderately ill, acutely ill) that the individual observer is seeing. Also systematic error arising from the diagnostic process (which may be determined by the culture, customs, or individual idiosyncrasy of the person providing care for the patient).

BIAS, IN ASSUMPTION (Syn: conceptual bias) Error arising from faulty logic or premises or mistaken beliefs on the part of the investigator. False conclusions about the explanation for associations between variables. Example: Having correctly deduced the mode of transmission of cholera, John Snow concluded that yellow fever was transmitted by similar means. In fact, the "miasma" theory would better fit the facts of yellow fever transmission.

BIAS IN AUTOPSY SERIES Systematic error resulting from the fact that autopsies represent a nonrandom sample of all deaths.

BIAS, BERKSON'S See **BIAS, SELECTION**.

BIAS DUE TO CONFOUNDING See **CONFOUNDING**.

BIAS, DESIGN The difference between a true value and that actually obtained, occurring as a result of faulty design of a study. Some examples are (1) uncontrolled studies where the effects of two processes cannot be separated (confounding), (2) controlled studies where observations are based on a poorly defined population, and (3) nonsimultaneous comparisons, e.g., use of historical controls.

BIAS, DETECTION Due to systematic error(s) in methods of ascertainment, diagnosis, or verification of cases in an epidemiologic survey, study, or investigation. Example: Verification of diagnosis by laboratory tests in hospital cases, but failure to apply the same tests to cases outside the hospital.

BIAS DUE TO DIGIT PREFERENCE See **DIGIT PREFERENCE**.

BIAS IN HANDLING OUTLIERS Error arising from a failure to discard an unusual value occurring in a small sample, or due to exclusion of unusual values that should be included.

BIAS, INFORMATION (Syn: observational bias) A flaw in measuring exposure or outcome that results in differential quality (accuracy) of information between compared groups.

BIAS DUE TO INSTRUMENTAL ERROR Systematic error due to faulty calibration, inaccurate measuring instruments, contaminated reagents, incorrect dilution or mixing of reagents, etc.

BIAS OF INTERPRETATION Error arising from inference and speculation. Sources of the

error include (1) failure of the investigator to consider every interpretation consistent with the facts and to assess the credentials of each, and (2) mishandling of cases that constitute exceptions to some general conclusion.

BIAS, INTERVIEWER Systematic error due to interviewers' subconscious or even conscious gathering of selective data.

BIAS, "LEAD-TIME" A systematic error arising when follow-up of two groups does not begin at strictly comparable times. Occurs especially when one group has been diagnosed earlier in the natural history of the disease than the other group. See also **ZERO TIME SHIFT**.

BIAS, LENGTH A systematic error due to the selection of a disproportionate number of long-duration cases (cases who survive longest) in one group and not in the other. Can occur when prevalent cases, rather than incident cases, are included in a case control study.

BIAS, MEASUREMENT Systematic error arising from inaccurate measurement (or classification) of subjects on the study variables.

BIAS, OBSERVER Systematic difference between a true value and that actually observed due to observer variation. Observer variation may be due to differences among observers (interobserver variation) or to variation in readings by the same observer on separate occasions (intraobserver variation). See also **OBSERVER VARIATION**.

BIAS IN THE PRESENTATION OF DATA Error due to irregularities produced by **DIGIT PREFERENCE**, incomplete data, poor techniques of measurement, or technically poor laboratory standards.

BIAS IN PUBLICATION An editorial predilection for publishing particular findings, e.g., positive results, which leads to the failure of authors to submit negative findings for publication or failure of journal editors to accept and publish reports with negative findings. This can distort the general belief about what has been demonstrated in a particular situation.

BIAS OF AN ESTIMATOR The difference between the expected value of an estimator of a parameter and the true value of this parameter. See also **UNBIASED ESTIMATOR**.

BIAS, RECALL Systematic error due to differences in accuracy or completeness of recall to memory of prior events or experiences. Example: Mothers whose children have had or have died of leukemia are more likely than mothers of healthy living children to remember details of diagnostic x-ray examinations to which these children were exposed in utero.

BIAS, REPORTING Selective suppression or revealing of information such as past history of sexually transmitted disease.

BIAS, RESPONSE Systematic error due to difference in characteristics between those who choose or volunteer to participate in a study and those who do not.

BIAS, SAMPLING Unless the sampling method ensures that all members of the "universe" or reference population have a known chance of selection in the sample, bias is possible. The best way to ensure a known chance of selection for all is to use a probability sampling method such as a table of random numbers.

BIAS, SELECTION Error due to systematic differences in characteristics between those who are selected for study and those who are not. Examples include hospital cases or cases under a physician's care, excluding those who die before admission to hospital because the course of their disease is so acute, those not sick enough to require hospital care, or those excluded by distance, cost, or other factors. Selection bias also invalidates generalizable conclusions from surveys that would include only volunteers from a healthy population.

A special example is **BERKSON'S BIAS**,¹ which Berkson characterized as the set of

selective factors that lead hospital cases and controls in a case control study to be systematically different from one another. This occurs when the combination of exposure and disease under study increases the risk of hospital admission, thus leading to a systematically higher exposure rate among the hospital cases than the hospital controls. This in turn results in systematic distortion of the ODDS RATIO.

¹Berkson J: Limitations of the application of fourfold table analysis to hospital data. *Biometrics Bull* 2:47-53, 1946.

BIAS DUE TO WITHDRAWALS A difference between the true value and that actually observed in a study due to the characteristics of those subjects who choose to withdraw.

BILLS OF MORTALITY Weekly and annual abstracts of christenings and burials, distinguishing deaths from the plague, compiled for London (and some other cities), especially in times of plague, from the English parish registers that started in 1538. From 1629, the annual bill was published regularly and included a breakdown of deaths by cause. These records were the basis for the earliest vital statistics, compiled, analyzed, and discussed by John Graunt in *Natural and Political Observations . . . on the Bills of Mortality* (1662).

BIMODAL DISTRIBUTION A distribution with two regions of high frequency separated by a region of low frequency of observations. A two-peak distribution.

BINARY VARIABLE A variable having only two possible values, e.g. on or off, 0 or 1. See also BIT.

BINOMIAL DISTRIBUTION A probability distribution associated with two mutually exclusive outcomes, e.g., presence or absence of a clinical or laboratory sign, death, or survival. The probability distribution of the number of occurrences of a binary event in a sample of n independent observations. The binomial distribution is used to model CUMULATIVE INCIDENCE RATES and PREVALENCE RATES. The BERNOULLI DISTRIBUTION is a special case of the binomial distribution with $n = 1$.

BIOASSAY The quantitative evaluation of the potency of a substance by assessing its effects on tissues, cells, live experimental animals, or humans.

Bioassay may be a direct method of estimating relative potency: groups of subjects are assigned to each of two (or more) preparations; the dose that is just sufficient to produce a specified response is measured, and the estimate is the ratio of the mean doses for the two (or more) groups. In this method, the death of the subject may be used as the "response."

The indirect method (more commonly used) requires study of the relationship between the magnitude of a dose and the magnitude of a quantitative response produced by it.

BIOLOGICAL PLAUSIBILITY The criterion that an observed, presumably or putatively causal ASSOCIATION fits previously existing biological or medical knowledge. This judgment should be used cautiously since it could impede development of new knowledge that does not fit existing ideas.

BIOLOGICAL TRANSMISSION See VECTOR-BORNE INFECTION.

BIOMETRY [literally, the measurement of life] The application of statistical methods to the study of numerical data based on biological observations and phenomena. The term was coined by W. F. R. Weldon (1860-1906), a zoologist at University College, London. FRANCIS GALTON has been called "the father of biometry" for his application of statistical methods to the analysis of biological variation. However, others preceded him, e.g., QUETELET and LOUIS.

BIOSTATISTICS Application of STATISTICS to biological problems. The term is considered

2503012397

by many biomedical scientists to mean the application of statistics specifically to medical problems, but its real meaning is broader.

BIRAUD, YVES (1900-1965) French physician and statistician. He served the League of Nations and later WHO as Director of Epidemiological and Statistical Services from 1925 to 1960. In 1960, he founded the first chair of Health Statistics in France, at the *Ecole de santé publique* in Rennes.

BIRTH CERTIFICATE Official, legal document recording details of a live birth, usually comprising name, date, place, identity of parents, and sometimes additional information such as birth weight. It provides the basis for vital statistics of birth and birthrates in a political or administrative jurisdiction, and for the denominator for infant mortality and certain other vital rates.

BIRTH COHORT See COHORT.

BIRTH COHORT ANALYSIS See COHORT ANALYSIS.

BIRTH INTERVAL Interval between termination of one completed pregnancy and the termination of the next.

BIRTH ORDER The ranking of siblings according to age, starting with the eldest in a family. The ordinal number of a given live birth in relation to all previous live births of the same woman. Thus, 4 is the birth order of the fourth live birth occurring to the same woman. This strict demographic definition may be loosened to include all births, i.e., still-births as well as live births

BIRTH RATE A summary rate based on the number of live births in a population over a given period, usually one year.

$$\text{Birth rate} = \frac{\text{Number of live births to residents in an area in a calendar year}}{\text{Average or midyear population in the area in that year}} \times 1000$$

BIRTH WEIGHT Infant's weight recorded at the time of birth and, in some countries, entered on the birth certificate. Certain variants of birth weight are precisely defined. Low birth weight (LBW) is below 2500 g. Very low birth weight (VLBW) is below 1500 g. Ultralow birth weight (ULBW) is below 1000 g. Large for gestational age (LGA) is birth weight above the 90th percentile. Average weight for gestational age (AGA) (Syn: appropriate or adequate): birth weight between 10th and 90th percentiles. Small for gestational age (SGA) (Syn: small for dates): birth weight below 10th percentile.

BIT Acronym for binary digit; the signal in computing. See also BYTE.

"BLACK BOX" A jargon term, meaning a method of reasoning or studying a problem, in which the methods, procedures, etc., as such are not described, explained, or perhaps even understood. Nothing is stated or inferred about the method; discussion and conclusions relate solely to the empirical relationships observed. An alternative definition is the following: A method of formally relating an input, e.g., quantity of a drug absorbed over a period or a putative causal factor, to an output, e.g., the amount of the drug eliminated in a given period, or an observed effect, without making detailed assumptions about the mechanisms that have contributed to the transformation of input to output within the organism (the "black box").

BLIND(ED) STUDY (Syn: masked study) A study in which observer(s) and/or subjects are kept ignorant of the group to which the subjects are assigned, as in an experiment,

or of the population from which the subjects come, as in a nonexperimental study. When both observer and subjects are kept ignorant, we refer to a double-blind study. If the statistical analysis is also done in ignorance of the group to which subjects belong, the study is sometimes described as triple-blind. The intent of keeping subjects and/or investigators blinded, i.e., unaware of knowledge that might introduce a bias, is to eliminate the effects of such biases. To avoid confusion about the meaning of the word "blind" some authors prefer to describe such studies as "masked."

BLOCKED RANDOMIZATION See STRATIFIED RANDOMIZATION. The analogue in a randomized experiment of individual matching in an observational study.

BODY MASS INDEX (Syn: Quetelet's index) One of the anthropometric measures of body mass. Defined as (weight) + (height)². This measure has the highest correlation with skinfold thickness or body density and in this respect is superior to the PONDERAL INDEX.

BOOTSTRAP A technique for estimating the variance and the bias of an estimator by repeatedly drawing random samples with replacement from the observations at hand. One applies the estimator to each sample drawn, thus obtaining a set of estimates. The observed variance of this set is the bootstrap estimate of variance. The difference between the average of the set of estimates and the original estimate is the bootstrap estimate of bias.

BREAKPOINT In helminth epidemiology, the critical mean wormload in a community, below which the helminth mating frequency is too low to maintain reproduction. A value exceeding the breakpoint of a wormload means that the wormload will increase until equilibrium is reached; a value less than or equal to the breakpoint means that the wormload will decrease progressively.

BYTE A group of adjacent bits, commonly 4, 6, or 8, operating as a unit for storage and manipulation of data in a computer. See also BIT.

2503012398

C

CALIPER MATCHING See MATCHING.

CANADIAN MORTALITY DATA BASE A large set of computer-stored death statistics; personal identifiers and causes of all deaths in Canada since 1950 have been computer-stored, and the death certificates have been preserved on microfiche. This data base and record linkage have been used in some important historical cohort studies. See also NATIONAL DEATH INDEX.

CANCER REGISTRY See REGISTER.

CARRIER

- I. A person or animal that harbors a specific infectious agent in the absence of discernible clinical disease and serves as a potential source of infection. The carrier state may occur in an individual with an infection that is inapparent throughout its course (known as healthy or asymptomatic carrier), or during the incubation period, convalescence, and postconvalescence of an individual with a clinically recognizable disease (known as incubatory carrier or convalescent carrier). The carrier state may be of short or long duration (temporary or transient carrier or chronic carrier).¹

¹Adapted from *Control of Communicable Disease in Man*, 14th ed. Washington, DC: American Public Health Association, 1985.

CARRYING CAPACITY An estimate of the numbers of people that a nation, region, or the planet can sustain.

CASE In epidemiology, a person in the population or study group identified as having the particular disease, health disorder, or condition under investigation. A variety of criteria may be used to identify cases, e.g., individual physicians' diagnoses, registries and notifications, abstracts of clinical records, surveys of the general population, population screening, and reporting of defects such as in a dental record. The epidemiologic definition of a case is not necessarily the same as the ordinary clinical definition.

CASE-BASE STUDY A study that starts with the identification and sampling of persons with the disease of interest, and then samples the entire base population (of cases and noncases) from which the original cases arose. This design is similar to a CASE CONTROL STUDY in most respects, but cases may appear in the comparison (base) sample as well as in the case sample.

CASE, COLLATERAL A case occurring in the immediate vicinity of a case which has been the subject of an epidemiological investigation; a term used mainly in malaria control programs, equivalent to the term contact as used in infectious disease epidemiology.

CASE COMPARISON STUDY See CASE CONTROL STUDY.

CASE COMPEER STUDY See CASE CONTROL STUDY.

CASE CONTROL STUDY (Syn: case comparison study, case compare study, case history study, case referent study, retrospective study) A study that starts with the identification of persons with the disease (or other outcome variable) of interest, and a suitable control (comparison, reference) group of persons without the disease. The relationship of an attribute to the disease is examined by comparing the diseased and nondiseased with regard to how frequently the attribute is present or, if quantitative, the levels of the attribute, in each of the groups.

Such a study can be called "retrospective" because it starts after the onset of disease and looks back to the postulated causal factors. Cases and controls in a case control study may be accumulated "prospectively;" that is, as each new case is diagnosed it is entered in the study. Nevertheless, such a study may still be called "retrospective" because it looks back from the outcome to its causes. The terms "cases" and "controls" are sometimes used to describe subjects in a RANDOMIZED CONTROLLED TRIAL but, the term "case control study" should not be used to describe such a study.

The terms "case control study" and "retrospective study" have been used most often to describe this method. Other terms also used are listed above. The concept of the case-control study is to be found in the works of P.C.A. Louis;¹ the first explicit description of the method is contained in a paper by William Augustus Guy, who reported his analysis of the relationship between prior occupational exposure and the occurrence of pulmonary consumption to the Statistical Society of London in 1843.² The evolution of the case-control study thereafter has been described by Lilienfeld and Lilienfeld.³ The first modern use of the method was a case-control study of breast cancer, reported by Lane-Claypon⁴ in 1926; from that time onward, case-control studies became increasingly popular and widely used.

¹Louis PCA: Researches on Phthisis; Anatomical, Pathological and Therapeutical. (Trans. W.H. Wolse). London: Sydenham Society, 1844.

²Guy WA: Contributions to a knowledge of the influence of employments on health. *J Roy Stat Soc* 6:197-211, 1843.

³Lilienfeld AM, Lilienfeld D: A century of case-control studies—progress. *J Chron Dis* 32:5-19, 1979.

⁴Lane-Claypon JE: A further report on cancer of the breast. *Rpt Pub Hlth Med Subj* 32. London: HMSO, 1926.

CASE FATALITY RATE The proportion of cases of a specified condition which are fatal within a specified time.

$$\text{Case fatality rate (usually expressed as a percentage)} = \frac{\text{Number of deaths from a disease (in a given period)}}{\text{Number of diagnosed cases of that disease (in the same period)}} \times 100$$

This definition can lead to paradox when more persons die of the disease than develop it during a given period. For instance, chemical poisoning that is slowly but inexorably fatal may cause many persons to develop the disease over a relatively short period of time, but the deaths may not occur until some years later and may be spread over a period of years during which there are no new cases. Thus, in calculating the case fatality rate, it is necessary to acknowledge that the time dimension varies: it may be brief, e.g., covering only the period of stay in a hospital, of finite duration, e.g., one year, or of longer duration still. The term "case fatality rate" is then better replaced by a term such as "survival rate" or by the use of a SURVIVORSHIP TABLE. See also ATTACK RATE.

2503012399

CASE HISTORY STUDY

1. Synonym for CASE CONTROL STUDY.
2. In clinical medicine, a case report, or a report on a series of cases.

CASE REFERENT STUDY See CASE CONTROL STUDY.

CATASTROPHE THEORY A branch of mathematics dealing with large changes in the total system that may result from small changes in a critical variable in the system. An example is the sudden change in the physical state of water into steam or ice with rise or fall of temperature beyond a critical level. Certain epidemics, gene frequencies, and behavioral phenomena in populations may abide by the same mathematical rule. Herd immunity is an example.

CATCHMENT AREA Region, which may be well- or ill-defined, from which the clients of a particular health facility are drawn.

CAUSALITY The relating of causes to the effects they produce. Most of epidemiology concerns causality and several types of causes can be distinguished. It should be clearly stated, however, that epidemiologic evidence by itself is insufficient to establish causality.

A cause is termed "necessary" when it must always precede an effect. This effect need not be the sole result of the one cause. A cause is termed "sufficient" when it inevitably initiates or produces an effect. Any given cause may be necessary, sufficient, neither, or both. These possibilities are explained below.

Four conditions under which independent variable X may cause Y

variable X may cause Y

X is necessary X is sufficient

| | X is necessary | X is sufficient |
|----|----------------|-----------------|
| 1. | + | + |
| 2. | + | - |
| 3. | - | + |
| 4. | - | - |

1. X is necessary and sufficient to cause Y. Both X and Y are always present together, and nothing but X is needed to cause Y; X→Y.

2. X is necessary but not sufficient to cause Y. X must be present when Y is present, but Y is not always present when X is. Some additional factor(s) must also be present; X and Z→Y.

3. X is not necessary but is sufficient to cause Y. Y is present when X is, but X may or may not be present when Y is present, because Y has other causes and can occur without X. For example, an enlarged spleen can have many separate causes that are unconnected with each other; X→Y; Z→Y.

4. X is neither necessary nor sufficient to cause Y. Again, X may or may not be present when Y is present. Under these conditions, however, if X is present with Y, some additional factor must also be present. Here X is a contributory cause of Y in some causal sequences; X and Z→Y; W and Z→Y. These relationships and the logic of causal inference are discussed in *Causal Inference*.¹

¹Rothman KJ (Ed): *Causal Inference*. Chestnut Hill, MA: Epidemiology Resources Inc., 1988.

CAUSATION OF DISEASE, FACTORS IN The following factors have been differentiated (but they are not mutually exclusive):

Predisposing factors are those that prepare, sensitize, condition, or otherwise create a situation such as a level of immunity or state of susceptibility so that the host tends to react in a specific fashion to a disease agent, personal interaction, environmental stimulus, or specific incentive. Examples include age, sex, marital status,

family size, educational level, previous illness experience, presence of concurrent illness, dependency, working environment, and attitudes toward the use of health services. These factors may be "necessary" but are rarely "sufficient" to cause the phenomenon under study.

Enabling factors are those that facilitate the manifestation of disease, disability, ill-health, or the use of services or conversely those that facilitate recovery from illness, maintenance or enhancement of health status, or more appropriate use of health services. Examples include income, health insurance coverage, nutrition, climate, housing, personal support systems, and availability of medical care. These factors may be "necessary" but are rarely "sufficient" to cause the phenomenon under study.

Precipitating factors are those associated with the definitive onset of a disease, illness, accident, behavioral response, or course of action. Usually one factor is more important or more obviously recognizable than others if several are involved and one may often be regarded as "necessary." Examples include exposure to specific disease, amount or level of an infectious organism, drug, noxious agent, physical trauma, personal interaction, occupational stimulus, or new awareness or knowledge.

Reinforcing factors are those tending to perpetuate or aggravate the presence of a disease, disability, impairment, attitude, pattern of behavior, or course of action. They may tend to be repetitive, recurrent, or persistent and may or may not necessarily be the same or similar to those categorized as predisposing, enabling, or precipitating. Examples include repeated exposure to the same noxious stimulus (in the absence of an appropriate immune response) such as an infectious agent, work, household, or interpersonal environment, presence of financial incentive or disincentive, personal satisfaction, or deprivation.

CAUSES OF DEATH See DEATH CERTIFICATE.

CAUSE-DELETED LIFE TABLE A life table constructed using death rates lowered by eliminating the risk of dying from a specified cause; its most common use is to calculate the gain in life expectancy that would result from the elimination of one cause.

CAUSE-SPECIFIC RATE A rate that specifies events, such as deaths, according to their cause.

CENSORING This term refers to the loss of subjects from a follow-up study; the occurrence of the event of interest among such subjects is uncertain after a specified time when it was known that the event of interest had *not* occurred; it is not known, however, if or when the event of interest occurred subsequently. Such subjects are described as censored. For example, in a follow-up study with myocardial infarction as the outcome of interest, a subject who has not had an infarct but is killed in a traffic crash in year 6 is described as censored as of year 6, since it cannot be known when, if ever, he might have had an infarct at a later year of follow-up. This is censoring by competing risk; other varieties include loss to follow-up and termination of the study. Examination of data for censoring requires the use of special analytic methods, such as life table analysis.

CENSUS An enumeration of a population, originally intended for purposes of taxation and military service. Census enumeration of a population usually records identities of all persons in every place of residence, with age, or birth date, sex, occupation, national origin, language, marital status, income, and relationship to head of household, in addition to information on the dwelling place. Many other items of information may be included, e.g., educational level (or literacy), and health-related data such as permanent disability. A *de facto* census allocates persons according to their location at the time of enumeration. A *de jure* census assigns persons according to their usual place of residence at the time of enumeration.

CENSUS TRACT An area for which details of population structure are separately tabulated at a periodic census; normally it is the smallest unit of analysis of (published) census tabulations. Census tracts are chosen because they have well-defined boundaries, sometimes the same as local political jurisdictions, sometimes defined by conspicuous geographical features such as main roads, rivers. In urban areas census tracts may be further subdivided, e.g., into city blocks, but published tables do not contain details to this level.

CENTILE SEE QUANTILES.

CESSATION EXPERIMENT Controlled study in which an attempt is made to evaluate the termination of an exposure to risk such as a living habit that is considered to be of etiologic importance.

CHART The medical dossier of a patient. See also INFORMATION SYSTEM; MEDICAL RECORD.

CHECK DIGIT A single digit, derived from a multidigit number such as a case identification number, that is used as a screening test for transcription errors.

CHEMOPROPHYLAXIS The administration of a chemical, including antibiotics, to prevent the development of an infection or the progression of an infection to active manifest disease.

CHEMOTHERAPY The use of a chemical to treat a clinically recognizable disease or to limit its further progress.

CHILD DEATH RATE The number of deaths of children aged 1–4 years in a given year per 1000 children in this age group. This is a useful measure of the burden of preventable communicable diseases in the child population.

CHI-SQUARE (χ^2) DISTRIBUTION A variable is said to have a chi-square distribution with K degrees of freedom if it is distributed like the sum of the squares of K independent random variables, each of which has a normal distribution with mean zero and variance one.

CHI-SQUARE (χ^2) TEST Any statistical test based on comparison of a test statistic to a chi-square distribution. The oldest and most common chi-square tests are for detecting whether two or more population distributions differ from one another; these tests usually involve counts of data, and may involve comparison of samples from the distributions under study, or the comparison of a sample to a theoretically expected distribution. The Pearson chi-square test is probably the best known; another is the Mantel-Haenszel test. (Statisticians disagree about the terminal letter; a bare majority of those who contributed to the discussion of this entry prefer "chi-square" rather than "chi-squared." Either usage is acceptable.)

CHRISOMS This word, which appears in **BILLS OF MORTALITY**, means infants who die before formal baptism; therefore, the number recorded in Bills of Mortality can be used to estimate (albeit inaccurately) neonatal death rates in studies of historical demography and epidemiology.

CHRONIC 1. Referring to a health-related state, lasting a long time. 2. Referring to exposure, prolonged or long-term, often with specific reference to low-intensity. 3. The U.S. National Center for Health Statistics defines a "chronic" condition as one of three months' duration or longer.

CLASS A term used in the theory of frequency distributions. The total number of observations made upon a particular variate may be grouped into classes according to convenient divisions of the variate range in order to make subsequent analyses less laborious, or for other reasons. A group so determined is called a "class." The variate values that determine the upper and lower limits of a class are called "class boundaries," the interval between them is the class interval, and the frequency fall-

CLASSIFICATION (Syn: categorization) Assignment to predesignated classes on the basis of perceived common characteristics. A means of giving order to a group of disconnected facts. Ideally, a classification should be characterized by (1) naturalness—the classes correspond to the nature of the thing being classified, (2) exhaustiveness—every member of the group will fit into one (and only one) class in the system, (3) usefulness—the classification is practical, (4) simplicity—the subclasses are not excessive, and (5) constructability—the set of classes can be constructed by a demonstrably systematic procedure.

CLASSIFICATION OF DISEASES Arrangement of diseases into groups having common characteristics. Useful in efforts to achieve standardization, and therefore comparability, in the methods of presentation of mortality and morbidity data from different sources. May include a systematic numerical notation for each disease entry.

Examples include the **INTERNATIONAL CLASSIFICATION OF DISEASES, INJURIES, AND CAUSES OF DEATH (ICD)** and the **INTERNATIONAL CLASSIFICATION OF HEALTH PROBLEMS IN PRIMARY CARE (ICHPPC)**.

CLASS, SOCIAL A method of socially stratifying populations, e.g., according to education, income, or occupation. See also **SOCIOECONOMIC CLASSIFICATION**.

CLINICAL DECISION ANALYSIS Application of **DECISION ANALYSIS** in a clinical setting with the aim of applying epidemiologic and other data on probability of outcomes when alternative decisions can be made, e.g., surgical intervention or drug treatment for myocardial ischemia.

CLINICAL EPIDEMIOLOGIST A practitioner of clinical epidemiology.

CLINICAL EPIDEMIOLOGY While some epidemiologists deplore any adjectival qualification of the discipline, a subspecialty of clinical epidemiology is sufficiently demarcated to justify definition. There are plenty of suggested definitions. John R. Paul¹ proposed "A marriage between quantitative concepts used by epidemiologists to study disease in populations and decision-making in the individual case which is the daily fare of clinical medicine." Patient care is central to Sackett's definition²: "The application, by a physician who provides direct patient care, of epidemiologic and biometric methods to the study of diagnostic and therapeutic processes in order to effect an improvement in health." While limiting the discipline to medical graduates in clinical practice, this definition is conceptually close to the definition of clinical decision analysis; the proper distinction between clinical epidemiology and clinical decision analysis may be that the epidemiologist works with a defined population, even if it is a population of patients rather than a community-based population with numerator and denominator in the conventional epidemiologic sense; clinical decision analysis can be applied to a single patient. Abramson's definition³ is "The use of epidemiological principles, methods and findings in personal health care or community-oriented primary care, with special reference to applications in diagnostic and prognostic appraisal, decisions concerning care and the evaluation of care. The term sometimes refers to any epidemiological study conducted in a clinical setting." Weiss⁴ defines clinical epidemiology as "The study of variation in the outcome of illness and of the reasons for that variation." The existence of the above and other subtly different definitions suggests that this branch of epidemiology remains inchoate.

¹J Clin Invest 17:539-541, 1938.

²Am J Epidemiol 89:125-128, 1969.

³Personal communication, 1986.

⁴Clinical Epidemiology. New York: Oxford University Press, 1986.

CLINICAL TRIAL (Syn: therapeutic trial) A research activity that involves the administration of a test regimen to humans to evaluate its efficacy and safety. The term is

subject to wide variation in usage, from the first use in humans without any control treatment to a rigorously designed and executed experiment involving test and control treatments and randomization.

See also **COMMUNITY TRIAL**.

CLINIMETRICS Feinstein,¹ who coined this term, defines it as the domain concerned with indexes, rating scales, and other expressions that are used to describe or measure symptoms, physical signs, and other distinctly clinical phenomena in clinical medicine. Such measurements, of course, are an essential part of many epidemiologic studies.

¹Feinstein AR: *Clinimetrics*. New Haven and London: Yale University Press, 1987.

CLOSED COHORT A population in which membership begins at a defined time or with a defined event and ends only through occurrence of the study outcome or the end of eligibility for membership. An example is a population of women in labor being studied to determine the vital status of their offspring (i.e., whether live or stillborn).

CLUSTER ANALYSIS A set of statistical methods used to group variables or observations into strongly interrelated subgroups.

CLUSTERING (Syn: disease cluster, time cluster, time-place cluster) A closely grouped series of events or cases of a disease or other health-related phenomena with well-defined distribution patterns, in relation to time or place or both. The term is normally used to describe aggregation of relatively uncommon events or diseases, e.g., leukemia, multiple sclerosis.

CLUSTER SAMPLING A sampling method in which each unit selected is a group of persons (all persons in a city block, a family, etc.) rather than an individual.

CODING Translation of information, e.g., questionnaire responses, into numbered categories for entry in a data processing system.

COEFFICIENT OF VARIATION The ratio of the standard deviation to the mean. This is meaningful only if the variable is measured on a ratio scale. See **MEASUREMENT SCALE**.

COHORT [from Latin *cohors*, warriors, the tenth part of a legion]

1. The component of the population born during a particular period and identified by period of birth so that its characteristics (e.g., causes of death and numbers still living) can be ascertained as it enters successive time and age periods.
2. The term "cohort" has broadened to describe any designated group of persons who are followed or traced over a period of time, as in **COHORT STUDY** (prospective study).

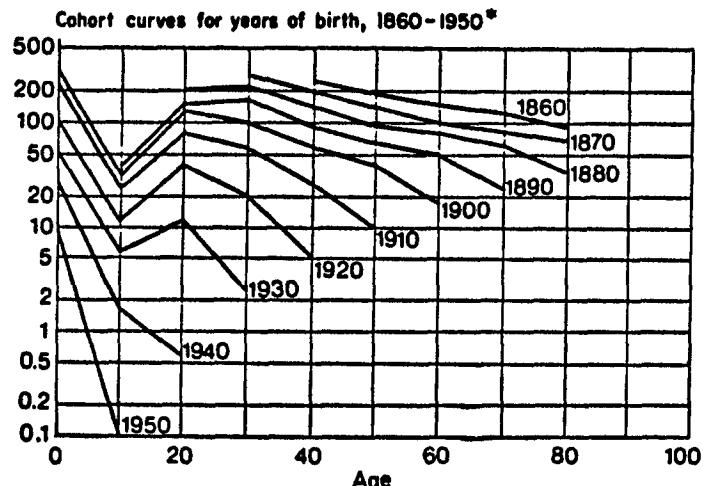
COHORT ANALYSIS The tabulation and analysis of morbidity or mortality rates in relationship to the ages of a specific group of people (cohort), identified at a particular period of time and followed as they pass through different ages during part or all of their life span. In certain circumstances, e.g., studies of migrant populations, cohort analysis may be performed according to duration of residence of migrants in a country rather than year of birth, in order to relate health or mortality experience to duration of exposure.

COHORT COMPONENT METHOD A method of population projection that takes the population distributed by age and sex at a base date and carries it forward in time on the basis of separate allowances for fertility, mortality, and migration.

COHORT EFFECT See **GENERATION EFFECT**.

COHORT INCIDENCE See **INCIDENCE**.

COHORT SLOPES Arrangement of data so that when plotted graphically, lines connect points representing the age-specific rates for population segments from the same



* The line associated with each year indicates death rates by age-group for persons born in that year

Cohort slopes (tuberculosis mortality rates of successive birth generations). Death rates for tuberculosis, by age, United States, 1900-1960 (per 100,000 population).

From Susser, Watson, Hopper, 1985.

generation of birth (see diagram). These slopes represent changes in rates with age during the life experience of each cohort.

COHORT STUDY (Syn: concurrent, follow-up, incidence, longitudinal, prospective study) The method of epidemiologic study in which subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed, or exposed in different degrees, to a factor or factors hypothesized to influence the probability of occurrence of a given disease or other outcome. The alternative terms for a cohort study, i.e., follow-up, longitudinal, and prospective study, describe an essential feature of the method, which is observation of the population for a sufficient number of person-years to generate reliable incidence or mortality rates in the population subsets. This generally implies study of a large population, study for a prolonged period (years), or both.

COINTERVENTION In a RANDOMIZED CONTROLLED TRIAL, the application of additional diagnostic or therapeutic procedures to members of either or both the experimental and the control groups.

COLD CHAIN A system of protection against high environmental temperatures for heat-labile vaccines, sera, and other active biological preparations. Unless the cold chain is preserved, such preparations are inactivated and immunization procedures, etc. will be ineffective. Preservation of the cold chain is an integral part of the WHO expanded program on immunization in tropical countries.

COLLINEARITY Very high correlation between variables.

COLONIZATION See INFECTION.

COMMENSAL Literally, eating together (sharing the same table); an organism that lives harmlessly in the gut. See also XENOBIOTIC.

COMMON SOURCE EPIDEMIC (Syn: common vehicle epidemic) See EPIDEMIC, COMMON SOURCE.

2503012402

COMMON VEHICLE SPREAD Spread of disease agent from a source that is common to those who acquire the disease, e.g., water, milk, shellfish, foods, air, or syringe contaminated by infectious or noxious agents. See also TRANSMISSION OF INFECTION.

COMMUNICABLE DISEASE (Syn: infectious disease) An illness due to a specific infectious agent or its toxic products that arises through transmission of that agent or its products from an infected person, animal, or reservoir to a susceptible host, either directly or indirectly through an intermediate plant or animal host, vector, or the inanimate environment. See also TRANSMISSION OF INFECTION.

COMMUNICABLE PERIOD The time during which an infectious agent may be transferred directly or indirectly from an infected person to another person, from an infected animal to man, or from an infected person to an animal, including arthropods. See also TRANSMISSION OF INFECTION.

COMMUNITY A group of individuals organized into a unit, or manifesting some unifying trait or common interest; loosely, the locality or catchment area population for which a service is provided, or more broadly, the state, nation, or body politic.

COMMUNITY DIAGNOSIS The process of appraising the health status of a community, including assembly of vital statistics and other health-related statistics and of information pertaining to determinants of health, such as prevalence of tobacco smoking, and examination of the relationships of these determinants to health in the specified community. The term may also denote the findings of this diagnostic process. Community diagnosis may attempt to be comprehensive, or may be restricted to specific health conditions, determinants, or subgroups. J.N. Morris¹ identified community diagnosis as one of the uses of epidemiology.

¹Br Med J 2:395-401, 1955.

COMMUNITY HEALTH See PUBLIC HEALTH.

COMMUNITY MEDICINE Since the late 1960s, this term has gained wide currency as the preferred name for important activities concerning health care in the community. There are several different definitions, including the following.

1. The field concerned with the study of health and disease in the population of a defined community or group. Its goal is to identify the health problems and needs of defined populations, to identify means by which these needs should be met, and to evaluate the extent to which health services effectively meet these needs.
2. The practice of medicine concerned with groups or populations rather than with individual patients. This includes the elements listed in definition 1, together with the organization and provision of health care at a community or group level.
3. The term is also used to describe the practice of medicine in the community, e.g., by a family physician. Some writers equate the terms "family medicine" and "community medicine"; others confine its use to public health practice.
4. Community-oriented primary health care is an integration of community medicine with the primary health care of individuals in the community. In this form of practice the community practitioner or community health team has responsibility for health care both at a community and at an individual level.

See also PUBLIC HEALTH; SOCIAL MEDICINE.

COMMUNITY TRIAL Experiment in which the unit of allocation to receive a preventive or therapeutic regimen is an entire community or political subdivision. Examples include the trials of fluoridation of drinking water, and of heart disease prevention in North Karelia (Finland) and California. See also CLINICAL TRIAL.

COMORBIDITY Disease(s) that coexist(s) in a study participant in addition to the index condition that is the subject of study.

COMPARISON GROUP Any group to which the index group is compared. Usually synonymous with control group.

COMPETING CAUSE When a previously common cause of death becomes rare, other causes become more prominent. These other causes are referred to as competing causes. For instance, among young adults, pneumonia and other infections were a common cause of death until about midway through the 20th century; their control has brought to prominence some competing causes of death, notably malignant disease and suicide.

COMPETING RISK An event that removes a subject from being at risk for the outcome under investigation. For example, in a study of smoking and cancer of the lung, a subject who dies of coronary heart disease is no longer at risk of lung cancer, and in this situation, coronary heart disease is a competing risk.

COMPLETED FERTILITY RATE The number of children born alive per woman in a cohort of women by the end of their child-bearing years.

COMPLETING THE CLINICAL PICTURE The use of epidemiology to define all modes of presentation of a disease, and/or all possible outcomes. One of the "uses of epidemiology" identified by J.N. Morris.¹

¹Br Med J 2:395-401, 1955.

COMPLETION RATE The proportion or percentage of persons in a SURVEY for whom complete data are available for analysis. See also RESPONSE RATE.

COMPOSITE INDEX An index, such as the Apgar score, Tumor/Nodes/Metastases (TNM) stage of cancer, that contains contributions from categories of several different variables.

COMPUTER A programmable electronic device that can be used to store and manipulate data in order to carry out designated functions. The two fundamental components of a computer are hardware, i.e., the actual electronic device, and software, the instructions or program used to carry out the function. Computer science has created a large language of its own, describing types of computers (main-frame, micro, digital, analogue, etc.) and all aspects of the process. Most of the terms used in this field are defined by AJ Meadows, M Gordon, and A Singleton.¹

¹Dictionary of New Information Technology. London: Century, 1982.

CONCORDANCE Pairs or groups of individuals of identical phenotype. In twin studies, a condition in which both twins exhibit or fail to exhibit a trait under investigation.

CONCORDANT A term used in TWIN STUDIES to describe a twin pair in which both twins exhibit a certain trait.

CONCURRENT STUDY See COHORT STUDY.

CONDITIONAL PROBABILITY The probability of an event, given that another event has occurred. If D and E are two events and $P(\cdot, \cdot)$ is "the probability of (\cdot, \cdot)", the conditional probability of D , given that E occurs, is denoted $P(D|E)$, where the vertical slash is read "given" and is equal to $P(D \text{ and } E)/P(E)$. The event E is the "conditioning event." Conditional probabilities obey all the axioms of probability theory. See also BAYES' THEOREM; PROBABILITY THEORY.

CONFIDENCE INTERVAL A range of values for a variable of interest, e.g., a rate, constructed so that this range has a specified probability of including the true value of the variable. The specified probability is called the confidence level, and the end points of the confidence interval are called the confidence limits.

CONFOUNDING [from the Latin *confundere*, to mix together]

1. A situation in which the effects of two processes are not separated. The dis-

tortion of the apparent effect of an exposure on risk brought about by the association with other factors that can influence the outcome.

2. A relationship between the effects of two or more causal factors as observed in a set of data, such that it is not logically possible to separate the contribution that any single causal factor has made to an effect.
3. A situation in which a measure of the effect of an exposure on risk is distorted because of the association of exposure with other factor(s) that influence the outcome under study.

CONFOUNDING VARIABLE (Syn: confounder) A variable that can cause or prevent the outcome of interest, is not an intermediate variable, and is not associated with the factor under investigation. Such a variable must be controlled in order to obtain an undistorted estimate of the effect of the study factor on risk.

CONSANGUINE Related by a common ancestor within the previous few generations.

CONSISTENCY

1. Close conformity between the findings in different samples, strata, or populations, or at different times or in different circumstances, or in studies conducted by different methods or different investigators. Consistency may be examined in order to study effect modification. Consistency of results on replication of studies is an important criterion in judgments of causality.
2. In statistics, an estimator is said to be consistent if the probability of it yielding estimates close to the true value approaches one as the sample size grows larger.

CONTACT (OF AN INFECTION) A person or animal that has been in such association with an infected person or animal or a contaminated environment as to have had opportunity to acquire the infection.

CONTACT, DIRECT A mode of transmission of infection between an infected host and susceptible host. Direct contact occurs when skin or mucous surfaces touch, as in shaking hands, kissing, and sexual intercourse. See also CONTAGION; TRANSMISSION OF INFECTION.

CONTACT, INDIRECT A mode of transmission of infection involving FOMITES OR VECTORS. Vectors may be mechanical (e.g., filth flies) or biological (the disease agent undergoes part of its life cycle in the vector species). See also TRANSMISSION OF INFECTION.

CONTACT, PRIMARY Person(s) in direct contact or associated with a communicable disease case.

CONTACT, SECONDARY Person(s) in contact or associated with a primary contact.

CONTAGION The transmission of infection by direct contact, droplet spread, or contaminated FOMITES. These are the modes of transmission specified by FRACASTORIUS in *De Contagione* (1546); contemporary usage is sometimes looser, but use of this term is best restricted to description of infection transmitted by direct contact.

CONTAGIOUS Transmitted by contact; in common usage, "highly infectious."

CONTAINMENT The concept of regional eradication of communicable disease, first proposed by Soper in 1949 for the elimination of smallpox.¹ Containment of a worldwide communicable disease demands a globally coordinated effort so that countries that have effected an interruption of transmission do not become reinfected following importation from neighboring endemic areas.

¹Pan American Health Organization, OSP, CE7, W-15, Washington DC, 1949.

CONTAMINATION

1. The presence of an infectious agent on a body surface; also on or in clothes, bedding, toys, surgical instruments or dressings, or other inanimate articles or substances including water, milk, and food. Pollution is distinct from contamination and implies the presence of offensive, but not necessarily infectious,

matter in the environment. Contamination of a body surface does not imply a carrier state. See also TRANSMISSION OF INFECTION.

2. The situation that exists when a population being studied for one condition or factor also possesses other conditions or factors that modify results of the study. In a RANDOMIZED CONTROLLED TRIAL, the inadvertent application of the experimental procedure to members of the control group, or inadvertent failure to apply the procedure to members of the experimental group.

CONTINGENCY TABLE A tabular cross-classification of data such that subcategories of one characteristic are indicated horizontally (in rows) and subcategories of another characteristic are indicated vertically (in columns). Tests of association between the characteristics in the columns and rows can be readily applied. The simplest contingency table is the fourfold, or 2×2 table. Contingency tables may be extended to include several dimensions of classification.

CONTINGENT VARIABLE See INTERMEDIATE VARIABLE.

CONTINUING SOURCE EPIDEMIC (OUTBREAK) An epidemic in which new cases of disease occur over a long period, indicating persistence of the disease source.

CONTINUOUS DATA, CONTINUOUS VARIABLE Data (variable) with a potentially infinite number of possible values along a continuum. Data representing a continuous variable include height, weight, and enzyme output.

CONTROL

1. (v.) To regulate, restrain, correct, restore to normal.
2. (n. or adj.) Applied to many communicable and some noncommunicable conditions, "control" means ongoing operations or programs aimed at reducing the incidence and/or prevalence, or eliminating such conditions.
3. (n.) As used in the expressions case-control study and randomized control(led) trial, "control" means person(s) in a comparison group that differs, respectively, in disease experience or allocation to a regimen, from the subjects of the study.
4. (v.) In statistics, "control" means to adjust for or take into account extraneous influences or observations.
5. (adj.) In the expression "control variable" we refer to an independent variable other than the hypothetical causal variable that has a potential effect on the dependent variable and is subject to control by analysis.

The use of the noun "control" to describe the comparison groups in a case control study and in a randomized control(led) trial can confuse the uninitiated, e.g., ethical review committees; the essential ethical distinction is that there may be no intervention in the lives or health status of the controls in a case-control study, whereas controls in a randomized controlled trial may be asked to undergo a procedure or regimen that may affect their health; their informed consent is therefore essential. Consent may not be required (save to gain access to medical records) to study controls in a case-control study. As M.W. Susser¹ has pointed out, the use of the word "control" as verb, adjective, and noun may confuse even careful readers. The verb is best used in the sense of controlling sources of extraneous variation in the dependent variable, whether by design or analysis. The verb is also used in the sense of controlling disease or its causes. The adjective is best used to describe control variables in contradistinction to uncontrolled and confounding variables. The adjective also can be used to describe a control group assembled for comparison with a group of cases or with an experimental group. The noun is best used to designate the members of a control group.

¹Causal Thinking in the Health Sciences. New York: Oxford, 1973.

CONTROLS, HISTORICAL Persons or patients used for comparison who had the condition

or treatment under study at a different time, generally at an earlier period than the study group or cases. Historical controls are often unsatisfactory because other factors affecting the condition under study may have changed to an unknown extent in the time elapsed.

CONTROLS, HOSPITAL Persons used for comparison who are drawn from the population of patients in a hospital. Hospital controls are often a source of SELECTION BIAS.

CONTROLS, MATCHED Controls who are selected so that they are similar to the study group, or cases, in specific characteristics. Some commonly used matching variables are age, sex, race, and socioeconomic status. See also MATCHING.

CONTROLS, NEIGHBORHOOD Persons used for comparison who live in the same locality as cases and therefore may resemble cases in environmental and socioeconomic criteria.

CONTROLS, SIBLING Persons used for comparison who are the siblings of cases and therefore share genetic makeup.

COORDINATES In a two-dimensional graph, the values of ordinate and abscissa that define the locus or position of a point.

CORDON SANITAIRE The barrier erected around a focus of infection. Used mainly in the isolation procedures applied to exclude cases and contacts of life-threatening communicable diseases from society. Mainly of historical interest.

CORRELATION The degree to which variables change together.

CORRELATION COEFFICIENT A measure of association that indicates the degree to which two variables have a linear relationship. This coefficient, represented by the letter r , can vary between +1 and -1; when $r = +1$, there is a perfect positive linear relationship in which one variable varies directly with the other; when $r = -1$, there is a perfect negative linear relationship between the variables. The measure can be generalized to quantify the degree of linear relationship between one variable and several others, in which case it is known as the multiple correlation coefficient. Kendall's Tau, Spearman's Rank Correlation, and Pearson's Product Moment Correlation tests are special varieties with occasional applications in epidemiology. M.G. Kendall and W.R. Buckland's *Dictionary of Statistical Terms*¹ gives details.

¹London: Longman, 1983.

CORRELATION, NONSENSE A meaningless correlation between two variables. Nonsense correlations sometimes occur when social, economic, or technological changes have the same trend over time as incidence or mortality rates. An example is correlation between the birth rate and the density of storks in parts of Holland and Germany. See also CONFOUNDING; ECOLOGICAL FALLACY.

COST-BENEFIT ANALYSIS An economic analysis in which the costs of medical care and the loss of net earnings due to death or disability are considered. The general rule for the allocation of funds in a cost-benefit analysis is that the ratio of marginal benefit (the benefit of preventing an additional case) to marginal cost (the cost of preventing an additional case) should be equal to or greater than 1.

COST-EFFECTIVENESS ANALYSIS This form of analysis seeks to determine the costs and effectiveness of an activity, or to compare similar alternative activities to determine the relative degree to which they will obtain the desired objectives or outcomes. The preferred action or alternative is one that requires the least cost to produce a given level of effectiveness, or provides the greatest effectiveness for a given level of cost. In the health care field, outcomes are measured in terms of health status.

COST-UTILITY ANALYSIS An economic analysis in which outcomes are measured in terms of their social value.

COVARIATE A variable that is possibly predictive of the outcome under study. A covariate may be of direct interest to the study or may be a confounding variable or effect modifier.

COVERAGE A measure of the extent to which the services rendered cover the potential need for these services in a community. It is expressed as a proportion in which the numerator is the number of services rendered, and the denominator is the number of instances in which the service should have been rendered. Example:

$$\text{Annual obstetric coverage} = \frac{\text{Number of deliveries attended by a qualified midwife or obstetrician}}{\text{Expected number of deliveries during the year in a given community}}$$

COX MODEL See PROPORTIONAL HAZARDS MODEL.

CRITERION A principle or standard by which something is judged. See also STANDARD.

CRONBACH'S ALPHA (Syn: internal consistency reliability) An estimate of the correlation between the total score across a series of items from a rating scale and the total score that would have been obtained had a comparable series of items been employed.

CROSS-CULTURAL STUDY A study in which populations from different cultural backgrounds are compared.

CROSSOVER DESIGN A method of comparing two or more treatments or interventions in which the subjects or patients, upon completion of the course of one treatment, are switched to another. In the case of two treatments, A and B, half the subjects are randomly allocated to receive these in the order A, B and half to receive them in the order B, A. A criticism of this design is that effects of the first treatment may carry over into the period when the second is given.

CROSS-PRODUCT RATIO See ODDS RATIO.

CROSS-SECTIONAL STUDY (Syn: disease frequency survey, prevalence study) A study that examines the relationship between diseases (or other health-related characteristics) and other variables of interest as they exist in a defined population at one particular time. The presence or absence of disease and the presence or absence of the other variables (or, if they are quantitative, their level) are determined in each member of the study population or in a representative sample at one particular time. The relationship between a variable and the disease can be examined (1) in terms of the prevalence of disease in different population subgroups defined according to the presence or absence (or level) of the variables and (2) in terms of the presence or absence (or level) of the variables in the diseased versus the nondiseased. Note that disease prevalence rather than incidence is normally recorded in a cross-sectional study. The temporal sequence of cause and effect cannot necessarily be determined in a cross-sectional study. See also MORBIDITY SURVEY.

CRUDE DEATH RATE See DEATH RATE.

CUMULATIVE DEATH RATE The proportion of a group that dies over a specified time interval. It may refer to all deaths or to deaths from specific cause(s). If follow-up is not complete on all persons the proper estimation of this rate requires the use of methods that take account of CENSORING. Distinct from FORCE OF MORTALITY.

CUMULATIVE INCIDENCE, CUMULATIVE INCIDENCE RATE The number or proportion of a group of people who experience the onset of a health-related event during a specified time interval; this interval is generally the same for all members of the group.

but, as in lifetime incidence, it may vary from person to person without reference to age.

CUMULATIVE INCIDENCE RATIO The ratio of the cumulative incidence rate in the exposed to the cumulative incidence rate in the unexposed.

CUSUM Acronym for cumulative sum (of a series of measurements). This is a useful way to demonstrate a change in trend or direction of a series of measurements.¹ Calculation begins with a reference figure, e.g. the expected average measurement. As each new measurement is observed, the reference figure is subtracted, and a cumulative total is produced by adding each successive difference. This cumulative total is the cusum.

¹Alderson M: An Introduction to Epidemiology, 2nd ed. London: Macmillan, 1983.

CYCPLICITY, SEASONAL The annual cycling of incidence on a seasonal basis. Certain acute infectious diseases, if of greater than rare occurrence, peak in one season of the year and reach the low point six months later (or in the opposite season). The onset of some symptoms of some chronic diseases also may show this amplitudinal cyclicity. Demographic phenomena such as marriage and births, and mortality from all causes and certain specific causes, may also exhibit seasonal cyclicity.

CYCPLICITY, SECULAR Long-term (greater than one year) cycling of disease incidence. For example, measles in a large, unimmunized population has a high incidence every second year; hepatitis A has a higher incidence every seventh year. Such cycling is the result of continuous exhaustion and replacement of susceptibles in a relatively stable population. Secular cyclicity may have large interval swings as in the recurrence of pandemics of influenza.

CYST COUNT See WORM COUNT.

D

DATA DREDGING A jargon term, meaning analyses done on a post hoc basis without benefit of pretested hypotheses, as a means of identifying noteworthy differences. Such analyses are sometimes done when data have been collected on a large number of variables and hypotheses are suggested by the data; the scientific validity of data dredging is at best dubious, usually unacceptable.

DATA PROCESSING Conversion (as by computer) of crude information into usable or storable form. Data generated by epidemiologic studies are usually transferred to punch cards or optical mark-sense forms and thence to a computer for storage and retrieval. The term is often loosely used to mean also the statistical analysis of data by a computer program. See also PUNCH CARD.

DEATH CERTIFICATE A vital record signed by a licensed physician or, in some nations, by another designated health worker, that includes cause of death, decedent's name, sex, birthdate, and place of residence and of death. Occupation, birthplace, and other information may be included. Immediate cause of death is recorded on the first line, followed by conditions giving rise to the immediate cause; the underlying cause is entered last. The underlying cause is coded and tabulated in official publications of cause-specific mortality. Other significant conditions may also be re-

| CAUSE OF DEATH | | Approximate interval between onset and death |
|--|--|--|
| I Disease or condition directly leading to death* | | |
| (a) due to (or as a consequence of) | | |
| (b) due to (or as a consequence of) | | |
| (c) | | |
| II Other significant conditions contributing to the death, but not related to the disease or condition causing it | | |
| | | |

* This does not mean the mode of dying, e.g., heart failure, asphyxia, etc. It means the disease, injury, or complication which caused death.

International Standard Death Certificate.

2503012406

corded separately, as is the mode of death, whether accidental or violent, etc. The most important entries on a death certificate are underlying causes of death and cause of death. These are defined in the Ninth (1975) Revision of the International Classification of Diseases, as follows:

Causes of death: The causes of death to be entered on the medical certificate of cause of death are all those diseases, morbid conditions, or injuries that either resulted in or contributed to death and the circumstances of the accident or violence which produced any such injuries.

Underlying cause of death: The underlying cause of death is (1) the disease or injury that initiated the train of events leading to death, or (2) the circumstances of the accident or violence that produced the fatal injury.

Personal identifying information such as birthplace, parents' names (last name at birth), and birthdates are included on death certificates in some jurisdictions; this extra information makes possible a range of RECORD LINKAGE studies.

DEATH RATE An estimate of the proportion of a population that dies during a specified period. The numerator is the number of persons dying during the period; the denominator is the size of the population, usually estimated as the mid-year population. The death rate in a population is generally calculated by the formula

$$\frac{\text{Number of deaths during a specified period}}{\text{Number of persons at risk of dying during the period}} \times 10^6$$

This rate is an estimate of the person-time death rate, i.e., the death rate per 10⁶ person-years. If the rate is low, it is also a good estimate of the cumulative death rate. This rate is also called the crude death rate.

DEATH REGISTRATION AREA A geographic area for which mortality data are published.

DECISION ANALYSIS A derivative of operations research and game theory that involves identifying all available choices and potential outcomes of each, in a series of decisions that have to be made about aspects of patient care—diagnostic procedures, therapeutic regimens, prognostic expectations. Epidemiologic data play a large part in determining the probabilities of outcomes following each choice that has to be made. The range of choices can be plotted on a decision tree, and at each branch, or decision node, the probabilities of each outcome that can be predicted are displayed. The decision tree thus portrays the choices available to those responsible for patient care and the probabilities of each outcome that will follow the choice of a particular action or strategy in patient care. The relative worth of each outcome is preferably also described as a utility or quality of life, e.g., a probability of life expectancy or of freedom from disability.¹

¹Pauker SG, Kassirer JP: Decision analysis. *N Engl J Med* 316:250-258, 1987.

DECISION TREE The alternative choices expressed in quantitative terms, available at each stage in the process of thinking through a problem, may be likened to branches, and the hierarchical sequence of options, to a tree. Hence, decision tree. It is a graphic device used in DECISION ANALYSIS, in which a series of decision options are represented as branches and subsequent possible outcomes are represented as further branches. The decisions and the eventualities are presented in the order they are likely to occur. The junction where a decision must be taken is called a decision node.

DEDUCTION Reasoned argument proceeding from the general to the particular.

DEGREES OF FREEDOM (df) The number of independent comparisons that can be made between the members of a sample. This important concept in statistical testing cannot be defined briefly. It refers to the number of independent contributions to a sampling distribution (such as χ^2 , t , and F distribution). In a CONTINGENCY TABLE it is one less than the number of row categories multiplied by one less than the number of column categories.

DEMAND (FOR HEALTH SERVICES) Willingness and/or ability to seek, use, and, in some settings, to pay for services. Sometimes further subdivided into *expressed demand* (equated with use) and *potential demand*, or NEED.

DEMOGRAPHIC TRANSITION The transition from high to low fertility and mortality rates, usually related to technological change and industrialization.

DEMOGRAPHY The study of populations, especially with reference to size and density, fertility, mortality, growth, age distribution, migration, and VITAL STATISTICS, and the interaction of all these with social and economic conditions.

DEMONSTRATION MODEL An experimental health care facility, program, or system with built-in provision for measuring aspects such as costs per unit of service, rates of use by patients or clients, and outcomes of encounters between providers and users. The aim usually is to determine the feasibility, efficacy, effectiveness, and/or efficiency of the model service.

DENOMINATOR The lower portion of a fraction used to calculate a rate or ratio. The population (or population experience, as in person-years, passenger-miles, etc.) at risk in the calculation of a rate or ratio. See also NUMERATOR.

DENSITY OF POPULATION Demographic term meaning numbers of persons in relation to available space.

DENSITY SAMPLING A method of selecting controls in a CASE CONTROL STUDY in which cases are sampled only from incident cases over a specific time period, and controls are sampled and interviewed throughout that period (rather than simply at one point in time, such as the end of the period). This method can reduce bias due to changing exposure patterns in the source population.

DEPENDENCY RATIO Proportion of children and old people in a population in comparison to all others, i.e., the proportion of economically inactive to economically active; "children" are usually defined as ages under 15 and "old people" as ages 65 and over.

DEPENDENT VARIABLE

1. A variable the value of which is dependent on the effect of other variable(s) [independent variable(s)] in the relationship under study. A manifestation or outcome whose variation we seek to explain or account for by the influence of independent variables.
2. In statistics, the dependent variable is the one predicted by a regression equation.

See also INDEPENDENT VARIABLE.

DESCRIPTIVE STUDY A study concerned with and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses. Contrast analytic study. An example is a community health survey, used to determine the health status of the people in a community. Descriptive studies, e.g., analyses of cancer registry data, can be used to measure risks.

DESIGN See RESEARCH DESIGN.

DESIGN VARIABLE

1. A study variable whose distribution in the subjects is determined by the investigator.

2. In statistics, a variable taking on the value 1 to indicate membership in a particular category and 0 or -1 to indicate nonmembership in the category. Used primarily in ANALYSIS OF VARIANCE.

DETERMINANT Any factor, whether event, characteristic, or other definable entity, that brings about change in a health condition, or other defined characteristic. See also CAUSALITY, FACTORS IN.

DIAGNOSIS The process of determining health status and the factors responsible for producing it; may be applied to an individual, family, group, or community. The term is applied both to the process of determination and to its findings. See also DISEASE LABEL.

DIAGNOSTIC INDEX A system for recording diagnoses, diseases, or problems of patients or clients in a medical practice or service, usually including identifying information (name, birthdate, sex) and dates of encounters. See also E-BOOK.

DIFFERENTIAL The difference(s) shown in tabulation of health and vital statistics according to age, sex, or some other factor; age differentials are the differences revealed in the tabulations of rates in age-groups, sex differentials are the differences in rates between males and females, income differentials are differences between designated income categories, etc.

DIGIT PREFERENCE A preference for certain numbers that leads to rounding off measurements. Rounding off may be to the nearest whole number, even number, multiple of 5 or 10, or (when time units like a week are involved) 7, 14, etc. This can be a form of OBSERVER VARIATION, or an attribute of respondent(s) in a survey.

DIMENSIONALITY The number of dimensions, i.e., scalar quantities, needed for accurate description of an element of a vector space.

DIRECT ADJUSTMENT, DIRECT STANDARDIZATION See STANDARDIZATION.

DISABILITY Temporary or long-term reduction of a person's capacity to function in society. See also INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES, AND HANDICAPS for the official WHO definition.

DISCORDANT A term used in TWIN STUDIES to describe a twin pair in which one twin exhibits a certain trait and the other does not. Also used in matched pair case control studies to describe a pair whose members had different exposures to the risk factor under study. Only the discordant pairs are informative about the association between exposure and disease.

DISCRETE DATA Data that can be arranged into naturally occurring or arbitrarily selected groups or sets of values, as opposed to data in which there are no naturally occurring breaks in continuity, i.e., CONTINUOUS DATA. An example is number of decayed, missing, and filled teeth (DMF).

DISCRIMINANT ANALYSIS A statistical analytic technique used with discrete dependent variables, concerned with separating sets of observed values and allocating new values; can sometimes be used instead of regression analysis. Kendall and Buckland¹ refer to this as "discriminatory analysis" and describe it as a rule for allocating individuals or values from two or more discrete populations to the correct population with minimal probability of misclassification.

¹Kendall MG, Buckland WR: *A Dictionary of Statistical Terms*, 4th ed. London: Longman, 1982.

DISEASE Literally, *dis-ease*, the opposite of *ease*, when something is wrong with a bodily function. The words "disease," "illness," and "sickness" are loosely interchangeable, but are better regarded as not wholly synonymous. M. W. Susser has suggested that they be used as follows:

Disease is a physiological/psychological dysfunction.

Illness is a subjective state of the person who feels aware of not being well;

Sickness is a state of social dysfunction, i.e., a role that the individual assumes when ill.

DISEASE FREQUENCY SURVEY See CROSS-SECTIONAL STUDY; MORBILITY SURVEY.

DISEASE LABEL The identity of the condition from which a patient suffers. It may be the name of a precisely defined disorder identified by a battery of tests, a probability statement based on consideration of what is most likely among several possibilities, or an opinion based on pattern recognition. Use of the word "label" can convey stigma, so this term should be used with care, if at all. See also DIAGNOSIS.

DISEASE ODDS RATIO See ODDS RATIO.

DISEASE, PRECLINICAL Disease with no signs or symptoms, because they have not yet developed. See also INAPPARENT INFECTION.

DISEASE REGISTRY See REGISTER, REGISTRY.

DISEASE, SUBCLINICAL A condition in which disease is detectable by special tests but does not reveal itself by signs or symptoms.

DISEASE TAXONOMY See TAXONOMY OF DISEASE.

DISINFECTION Killing of infectious agents outside the body by direct exposure to chemical or physical agents.

Concurrent disinfection is the application of disinfective measures as soon as possible after the discharge of infectious material from the body of an infected person, or after the soiling of articles with such infectious discharges, all personal contact with such discharges or articles being minimized prior to such disinfection.

Terminal disinfection is the application of disinfective measures after the patient has been removed by death or to a hospital, or has ceased to be a source of infection, or after other hospital isolation practices have been discontinued. Terminal disinfection is rarely practiced; terminal cleaning generally suffices, along with airing and sunning of rooms, furniture, and bedding. Disinfection is necessary only for diseases spread by indirect contact; steam sterilization or incineration of bedding and other items is desirable after a disease such as plague or anthrax.¹

¹Benenson AS (Ed): *Control of Communicable Diseases in Man*, 14th ed. Washington DC: American Public Health Association 1985.

DISINFESTATION Any physical or chemical process serving to destroy or remove undesired small animal forms, particularly arthropods or rodents, present upon the person, the clothing, or in the environment of an individual, or on domestic animals. Disinfestation includes delousing for infestation with *Pediculus humanus humanus*, the body louse. Synonyms include the terms "disinsection" and "disinsectization" when insects only are involved.

DISTRIBUTION The complete summary of the frequencies of the values or categories of a measurement made on a group of persons. The distribution tells either how many or what proportion of the group was found to have each value (or each range of values) out of all the possible values that the quantitative measure can have.

DISTRIBUTION-FREE METHOD A method which does not depend upon the form of the underlying distribution.

DISTRIBUTION FUNCTION A function that gives the relative frequency with which a random variable falls at or below each of a series of values. Examples include the normal distribution, log-normal distribution, chi-square distribution, *t* distribution, *F*-distribution, and binomial distribution, all of which have applications in epidemiology.

DMF The abbreviation DMF stands for decayed, missing, and filled teeth. Lowercase letters, i.e., dmf, are used for deciduous dentition, upper case for permanent teeth. The DMF number is widely used in dental epidemiology.

DOSE-RESPONSE RELATIONSHIP A relationship in which a change in amount, intensity, or duration of exposure is associated with a change—either an increase or a decrease—in risk of a specified outcome.

DOUBLE-BLIND TRIAL A procedure of blind assignment to study and control groups and blind assessment of outcome, designed to ensure that ascertainment of outcome is not biased by knowledge of the group to which an individual was assigned. "Double" refers to both parties, i.e., the observer(s) in contact with the subjects, and the subjects in the study and control groups. See also BLIND EXPERIMENT; RANDOMIZED CONTROLLED TRIAL.

DRIFT See GENETIC DRIFT; SOCIAL DRIFT.

DROPLET NUCLEI A type of particle implicated in the spread of airborne infection. Droplet nuclei are tiny particles (1–10 μm diameter) that represent the dried residue of droplets. They may be formed by (1) evaporation of droplets coughed or sneezed into the air or (2) aerosolization of infective materials. See also TRANSMISSION OF INFECTION.

DROPOUT A person enrolled in a study who becomes inaccessible or ineligible for follow-up, e.g., because of inability or unwillingness to remain enrolled in the study. The occurrence of dropouts can lead to biases in study results.

DUMMY VARIABLE See INDICATOR VARIABLE.

DYNAMIC POPULATION A population that gains and loses members; all natural populations are dynamic, a fact recognized by the term "population dynamics," used by demographers to denote changing composition. See also POPULATION DYNAMICS; STABLE POPULATION.

E

EARLY WARNING SYSTEM In disease surveillance, a specific procedure to detect as early as possible any departure from usual or normally observed frequency of phenomena. For example, the routine monitoring of numbers of deaths from pneumonia and influenza in large American cities is an early warning system for the identification of influenza epidemics. In developing countries, a change in children's average weights is an early warning signal of nutritional deficiency.

E-BOOK Method (developed by Eimerl)¹ of recording encounters in primary medical care: encounters are arranged by problem or diagnostic category, thus making it easy to count the number of persons seen (and the number of times each is seen) according to problem or diagnostic category in a given period of time. Widely used in epidemiologic studies of primary medical care. See also AGE-SEX REGISTER; DIAGNOSTIC INDEX.

¹Eimerl TS: Organized curiosity. *J Coll Gen Pract* 3:246-252, 1960.

ECOLOGICAL ANALYSIS Analysis based on aggregated or grouped data; errors in inference may result because associations may be artificially created or masked by the aggregation process.

ECOLOGICAL CORRELATION A correlation in which the units studied are populations rather than individuals. Correlations found in this manner may not hold true for the individual members of these populations. See also ECOLOGICAL FALLACY.

ECOLOGICAL FALLACY (Syn: aggregation bias, ecological bias)

1. The bias that may occur because an association observed between variables on an aggregate level does not necessarily represent the association that exists at an individual level.
2. An error in inference due to failure to distinguish between different levels of organization. A correlation between variables based on group (ecological) characteristics is not necessarily reproduced between variables based on individual characteristics; an association at one level may disappear at another, or even be reversed. Example: At the ecological level, a correlation has been found in several studies between the quality of drinking water and mortality rates from heart disease; it would be an ecological fallacy to infer from this alone that exposure to water of a particular level of hardness necessarily influences the individual's chances of getting or dying of heart disease.

ECOLOGICAL STUDY A study in which the units of analysis are populations or groups of people, rather than individuals. An example is the study of association between median income and cancer mortality rates in administrative jurisdictions such as states and counties.

ECOLOGY The study of the relationships among living organisms and their environment. "Human ecology" means the study of human groups as influenced by environmental factors, often including social and behavioral factors.

ECOSYSTEM The plant and animal life of a region considered in relation to the environmental factors that influence it; more specifically, the fundamental unit in ecology, comprising the living organisms and the nonliving elements that interact in a defined region.

EFFECT The result of a cause. In epidemiology, frequently a synonym for EFFECT MEASURE.

EFFECTIVENESS The extent to which a specific intervention, procedure, regimen, or service, when deployed in the field, does what it is intended to do for a defined population.

EFFECT MEASURE A quantity that measures the effect of a factor on the frequency or risk of a health outcome. Three such measures are attributable fractions, which measure the fraction of cases due to a factor; risk and rate differences, which measure the amount a factor adds to the risk or rate of a disease; and risk and rate ratios, which measure the amount by which a factor multiplies the risk or rate of disease.

EFFECT MODIFIER (Syn: conditional variable, moderator variable) A factor that modifies the effect of a putative causal factor under study. For example, age is an effect modifier for many conditions, and immunization status is an effect modifier for the consequences of exposure to pathogenic organisms. Effect modification is detected by varying the selected effect measure for the factor under study across levels of another factor. See also CAUSALITY, FACTORS IN; INTERACTION.

EFFECTIVE SAMPLE SIZE Sample size after dropouts, deaths, and other specified exclusions from an original sample.

EFFICACY The extent to which a specific intervention, procedure, regimen, or service produces a beneficial result under ideal conditions. Ideally, the determination of efficacy is based on the results of a RANDOMIZED CONTROLLED TRIAL.

EFFICIENCY

1. The effects or end-results achieved in relation to the effort expended in terms of money, resources, and time. The extent to which the resources used to provide a specific intervention, procedure, regimen, or service of known efficacy and effectiveness are minimized. A measure of the economy (or cost in resources) with which a procedure of known efficacy and effectiveness is carried out.
2. In statistics, the relative precision with which a particular study design or estimator will estimate a parameter of interest.

EGG COUNT See WORM COUNT.

ELIMINATION See ERADICATION (OF DISEASE).

EMPIRICAL Based directly on experience, e.g., observation or experiment, rather than on reasoning alone.

ENCOUNTER A face-to-face transaction between a personal health worker and a patient or client.

ENDEMIC DISEASE The constant presence of a disease or infectious agent within a given geographic area or population group; may also refer to the usual prevalence of a given disease within such area or group. See also HOLOENDEMIC DISEASE; HYPERENDEMIC DISEASE.

END RESULTS See OUTCOMES.

ENVIRONMENT All that which is external to the individual human host. Can be divided into physical, biological, social, cultural, etc., any or all of which can influence health status of populations.

EPIDEMIC [from the Greek *epi* (upon), *dēmos* (people)] The occurrence in a community or region of cases of an illness, specific health-related behavior, or other health-related events clearly in excess of normal expectancy. The community or region, and the period in which the cases occur, are specified precisely. The number of cases indicating the presence of an epidemic varies according to the agent, size, and type of population exposed, previous experience or lack of exposure to the disease, and time and place of occurrence; epidemicity is thus relative to usual frequency of the disease in the same area, among the specified population, at the same season of the year. A single case of a communicable disease long absent from a population or first invasion by a disease not previously recognized in that area requires immediate reporting and full field investigation; two cases of such a disease associated in time and place may be sufficient evidence to be considered an epidemic.

The word may be used also to describe outbreaks of disease in animal or plant populations. See also EPIZOOTIC; EPORNITHIC.

EPIDEMIC, COMMON SOURCE (Syn: common vehicle epidemic, holomiantic disease) Outbreak due to exposure of a group of persons to a noxious influence that is common to the individuals in the group. When the exposure is brief and essentially simultaneous, the resultant cases all develop within one incubation period of the disease (a "point" or "point source" epidemic).

The term "holomiantic disease" was used by Stallybrass (1931) to describe outbreaks of this type, but as with several other terms created from Greek or Latin roots, transmission to epidemiologists who lacked a classical education, did not take place.

EPIDEMIC CURVE A graphic plotting of the distribution of cases by time of onset.

EPIDEMIC, MATHEMATICAL MODEL OF See MATHEMATICAL MODEL.

EPIDEMIC, POINT SOURCE See EPIDEMIC, COMMON SOURCE.

EPIDEMIOLOGIST An investigator who studies the occurrence of disease or other health-related conditions or events in defined populations. The control of disease in populations is often also considered to be a task for the epidemiologist, especially in speaking of certain specialized fields such as malaria epidemiology. Epidemiologists may study disease in populations of animals and plants, as well as among human populations. See also CLINICAL EPIDEMIOLOGIST.

EPIDEMIOLOGY The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to control of health problems.

There have been many definitions of epidemiology. In the past 50 years or so, the definition has broadened from concern with communicable disease epidemics to take in all phenomena related to health in populations.

The *Oxford English Dictionary (OED)* gives as a definition: "That branch of medical science which treats of epidemics" and cites Parkin (1873) as a source. However, there was a "London Epidemiological Society" in the 1850s. The identity of the scholar who first used the word at that time has been lost. *Epidemiología* appears in the title of a Spanish history of epidemics, *Epidemiología española*, Madrid, 1802.

Epidemic is much older. The word appears in Johnson's *Dictionary* (1775), and *OED* gives a citation dated 1603. The word was, of course, used by Hippocrates.

EPIDEMIOLOGY, ANALYTIC See ANALYTIC STUDY.

EPIDEMIOLOGY, DESCRIPTIVE Study of the occurrence of disease or other health-related characteristics in human populations. General observations concerning the relationship of disease to basic characteristics such as age, sex, race, occupation, and social class; also concerned with geographic location. The major characteristics in descrip-

tive epidemiology can be classified under the headings: persons, place, and time. See also OBSERVATIONAL STUDY.

EPIDEMIOLOGY, EXPERIMENTAL See EXPERIMENTAL EPIDEMIOLOGY.

EPISODE Period in which a health problem or illness exists, from its onset to its resolution. See also ENCOUNTER.

EPIZOOTIC An outbreak (epidemic) of disease in an animal population (often with the implication that it may also affect human populations).

EPORNITHIC An outbreak (epidemic) of disease in a bird population.

ERADICATION (OF DISEASE) Termination of all transmission of infection by extermination of the infectious agent through surveillance and containment. Eradication, as in the instance of smallpox, was based on the joint activities of control and surveillance. Regional eradication has been successful with malaria and in some countries appears close to succeeding for measles. The term "elimination" is sometimes used to describe eradication of diseases such as measles from a large geographic region or political jurisdiction.

ERROR

1. A false or mistaken result obtained in a study or experiment. Several kinds of error can occur in epidemiology, for example, due to bias.
2. Random error is the portion of variation in a measurement that has no apparent connection to any other measurement or variable, generally regarded as due to chance.
3. Systematic error, which often has a recognizable source, e.g., a faulty measuring instrument, or pattern, e.g., it is consistently wrong in a particular direction. See also BIAS.

ERROR, TYPE I (Syn: alpha error) The error of rejecting a true null hypothesis. See also SIGNIFICANCE LEVEL; STATISTICAL TEST.

ERROR, TYPE II (Syn: beta error) The error of failing to reject a false null hypothesis. See also POWER; STATISTICAL TEST.

ESTIMATE A measurement or a statement about the value of some quantity is said to be an estimate if it is known, believed, or suspected to incorporate some degree of error.

ESTIMATOR In statistics, a function for computing estimates of a parameter from observed data.

ETHICS The branch of philosophy that deals with the distinction between right and wrong, with the moral consequences of human actions. Ethical principles govern the conduct of epidemiology, as they do all human activities; the ethical issues that are specific to epidemiological practice and research include informed consent, confidentiality, and respect for human rights. The issues have been defined, described, and discussed by many writers and by special committees under the auspices of research granting agencies and other official bodies in many countries.¹

¹See, for example, the following: Curran WJ: Protecting confidentiality in epidemiologic investigations by the Centers for Disease Control. *N Engl J Med* 314:1027-1028, 1986.

Susser MW, Stein Z, Kline J: Ethics in epidemiology. *Ann Amer Acad Pol Sci* 437:128-141, 1978. Commonwealth of Australia, National Health and Medical Research Council, Medical Research Ethics Committee: Report on Ethics in Epidemiological Research. Canberra, 1985.

Stolley PD: Faith, evidence and the epidemiologist. *J Public Health Pol* 6:37-42, 1985. Gordis, L, Gold E, Selser R: Privacy and protection in epidemiologic and medical research: Challenge and responsibility. *Am J Epidemiol* 105:163-168, 1977.

National Academy of Sciences, Institute of Medicine: *Ethics of Health Care*. Washington, DC, 1974. Tancredi LR (ed): *Ethical issues in epidemiologic research* (Vol. VII, series in Psychosocial Epidemiology). New Brunswick, NJ: Rutgers University Press, 1986.

ETHNIC GROUP A social group characterized by a distinctive social and cultural tradition, maintained within the group from generation to generation, a common history and origin, and a sense of identification with the group. Members of the group have distinctive features in their way of life, shared experiences, and often a common genetic heritage. These features may be reflected in their health and disease experience. See also RACE.

ETIOLOGY Literally, the science of causes, causality; in common usage, cause. See also CAUSALITY; PATHOGENESIS.

ETIOLOGIC FRACTION (EXPOSED) See ATTRIBUTABLE FRACTION (EXPOSED).

ETIOLOGIC FRACTION (POPULATION) See ATTRIBUTABLE FRACTION (POPULATION).

EVALUATION A process that attempts to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in the light of their objectives. Several varieties of evaluation can be distinguished, e.g., evaluation of structure, process, and outcome. See also CLINICAL TRIAL; EFFECTIVENESS; EFFICACY; EFFICIENCY; HEALTH SERVICES RESEARCH; PROGRAM EVALUATION AND REVIEW TECHNIQUES; QUALITY OF CARE.

EVAN'S POSTULATES Expanding biomedical knowledge has led to revision of HENLE'S and KOCH'S POSTULATES. Alfred Evans¹ developed those that follow, based on the Henle-Koch model.

1. Prevalence of the disease should be significantly higher in those exposed to the hypothesized cause than in controls not so exposed.
2. Exposure to the hypothesized cause should be more frequent among those with the disease than in controls without the disease—when all other risk factors are held constant.
3. Incidence of the disease should be significantly higher in those exposed to the hypothesized cause than in those not so exposed, as shown by prospective studies.
4. The disease should follow exposure to the hypothesized causative agent with a distribution of incubation periods on a bell-shaped curve.
5. A spectrum of host responses should follow exposure to the hypothesized agent along a logical biological gradient from mild to severe.
6. A measurable host response following exposure to the hypothesized cause should have a high probability of appearing in those lacking this before exposure (e.g., antibody, cancer cells), or should increase in magnitude if present before exposure. This response pattern should occur infrequently in persons not so exposed.
7. Experimental reproduction of the disease should occur more frequently in animals or man appropriately exposed to the hypothesized cause than in those not so exposed; this exposure may be deliberate in volunteers, experimentally induced in the laboratory, or may represent a regulation of natural exposure.
8. Elimination or modification of the hypothesized cause should decrease the incidence of the disease (i.e., attenuation of a virus, removal of tar from cigarettes).
9. Prevention or modification of the host's response on exposure to the hypothesized cause should decrease or eliminate the disease (i.e., immunization, drugs to lower cholesterol, specific lymphocyte transfer factor in cancer).
10. All of the relationships and findings should make biological and epidemiologic sense.

¹Evans AS: Causation and disease: The Henle-Koch postulates revisited. *Yale J Biol Med* 49:175-195, 1976.

EXACT METHOD A statistical method based on the actual, i.e., "exact" probability distribution of the study data, rather than on an approximation such as the normal or chi-square distribution; for example, Fisher's exact test.

EXACT TEST A statistical test based on the actual null probability distribution of the study data, rather than, say, normal approximation. The most common exact test is the Fisher-Irwin test for fourfold tables.

EXCESS RATE AMONG EXPOSED See RATE DIFFERENCE.

EXCESS RISK A term sometimes used to refer to the POPULATION EXCESS RATE and sometimes to RISK DIFFERENCE.

EXPANDED PROGRAMME ON IMMUNIZATION Part of the effort to achieve "Health for All by the Year 2000," under the auspices of WHO, UNICEF, and other international and bilateral aid agencies. This is a program of immunizing against diphtheria, tetanus, measles, pertussis, poliomyelitis, and tuberculosis, conducted especially in developing countries.

EXPECTATION OF LIFE (Syn: life expectancy or expectation) The average number of years an individual of a given age is expected to live if current mortality rates continue to apply. A statistical abstraction based on existing, age-specific death rates.

Life expectancy at birth (ℓ_0): Average number of years a newborn baby can be expected to live if current mortality trends continue. Corresponds to the total number of years a given birth cohort can be expected to live, divided by the number of children in the cohort. Life expectancy at birth is partly dependent on mortality in the first year of life and is lower in poor than in rich countries because of the higher infant and child mortality rates in the former.

Life expectancy at a given age, age x (ℓ_x): The average number of additional years a person age x would live if current mortality trends continue to apply, based on the age-specific death rates for a given year.

Life expectancy is a hypothetical measure and indicator of current health and mortality conditions. It is not a rate.

EXPERIMENT A study in which the investigator intentionally alters one or more factors under controlled conditions in order to study the effects of so doing.

EXPERIMENTAL EPIDEMIOLOGY In modern usage, this term is often equated with RANDOMIZED CONTROLLED TRIALS. To GREENWOOD and other epidemiologists in the 1920s, it meant the study of epidemics among colonies of experimental animals such as rats and mice. The original meaning of the term is preferable; if the word "experiment" is qualified by the adjective "epidemiologic" it is a synonym for RANDOMIZED CONTROLLED TRIAL. See also ANIMAL MODEL.

EXPERIMENTAL STUDY A study in which conditions are under the direct control of the investigator. In epidemiology, a study in which a population is selected for a planned trial of a regimen whose effects are measured by comparing the outcome of the regimen in the experimental group with the outcome of another regimen in a control group. To avoid bias members of the experimental and control groups should be comparable except in the regimen that is offered them. Allocation of individuals to experimental or control groups is ideally by randomization. In a RANDOMIZED CONTROLLED TRIAL, individuals are randomly allocated; in some experiments, e.g., fluoridation of drinking water, whole communities have been (nonrandomly) allocated to experimental and control groups.

EXPLANATORY STUDY A study whose main objective is to explain, rather than merely describe, a situation, by isolating the effects of specific variables and understanding the mechanisms of action. See also PRAGMATIC STUDY.

EXPLANATORY VARIABLE

1. A variable that causally explains the association or outcome under study.

2. In statistics, a synonym for INDEPENDENT VARIABLE.

EXPOSED In epidemiology, the exposed group (or simply, *the exposed*) is often used to connote a group whose members have been exposed to a supposed cause of a disease or health state of interest, or possess a characteristic that is a determinant of the health outcome of interest.

EXPOSURE

1. Proximity and/or contact with a source of a disease agent in such a manner that effective transmission of the agent or harmful effects of the agent may occur.
2. The amount of a factor to which a group or individual was exposed; sometimes contrasted with dose, the amount that enters or interacts with the organism.
3. Exposures may of course be beneficial rather than harmful, e.g., exposure to immunizing agents.

EXPOSURE-ODDS RATIO See ODDS RATIO.

EXPOSURE RATIO The ratio of rates at which persons in the case and control groups of a CASE CONTROL STUDY are exposed to the RISK FACTOR (or to the protective factor) of interest.

EXPRESSIVITY In genetics, the extent to which a gene is expressed.

EXTRAPOLATE, EXTRAPOLATION To predict the value of a variate outside the range of observations; the resulting prediction. See also INTERPOLATE.

EXTRINSIC INCUBATION PERIOD Time required for development of a disease agent in a vector from the time of uptake of the agent to the time when the vector is infective. See also INCUBATION PERIOD; VECTOR-BORNE INFECTION.

F

F DISTRIBUTION (Syn: Variance ratio distribution) The distribution of the ratio of two independent quantities each of which is distributed like a variance in normally distributed samples. So-named in honor of R.A. Fisher who first described this distribution.

F₁ ("F one") Term used in genetics to describe first-generation progeny of a mating.

FACTOR (Syn: determinant)

1. An event, characteristic, or other definable entity that brings about a change in a health condition or other defined outcome. See also CAUSALITY, CAUSATION OF DISEASE, FACTORS IN.
2. A synonym for (categorical) independent variable, or more precisely, an independent variable used to identify, with numerical codes, membership of qualitatively different groups. A causal role may be implied, as in "overcrowding is a factor in disease transmission" where overcrowding represents the highest level of the factor "crowding."

FACTOR ANALYSIS A set of statistical methods for analyzing the correlations among several variables in order to estimate the number of fundamental dimensions that underlie the observed data and to describe and measure those dimensions. Used frequently in the development of scoring systems for rating scales and questionnaires.

FACTORIAL DESIGN A method of setting up an experiment or study to assure that all levels of each intervention or classificatory factor occur with all levels of the others.

FALSE NEGATIVE Negative test result in a subject who possesses the attribute for which the test is conducted. The labeling of a diseased person as healthy when screening in the detection of disease. See also SCREENING; SENSITIVITY AND SPECIFICITY.

FALSE POSITIVE Positive test result in a subject who does not possess the attribute for which the test is conducted. The labeling of a healthy person as diseased when screening in the detection of disease. See also SCREENING; SENSITIVITY AND SPECIFICITY.

FAMILIAL DISEASE Disease that exhibits a tendency to familial occurrence. Familial occurrence of disease may be due to genetic transmission, intrafamilial transmission of infection or culture, interaction within the family, or the family's shared experience, including its exposure to a common environment.

FAMILY A group of two or more persons united by blood, adoptive or marital ties, or the common law equivalent; the family may include members who do not share the household but are united to other members by blood, adoptive or marital, or equivalent ties. Epidemiologic studies may be concerned with family members or with those who share the same household or dwelling unit.

FAMILY, EXTENDED A group of persons comprising members of several generations united by blood, adoptive and marital, or equivalent ties. See also FAMILY, NUCLEAR.

2503012412

FAMILY CONTACT DISEASE Disease that occurs among members of the family of a worker who is exposed to a toxic substance and carries this home on his person or his clothing, causing exposure to other family members.

FAMILY, NUCLEAR A group of persons comprising members of a single or at most two generations, usually husband-wife-children, united by blood or adoptive and marital or equivalent ties.

FAMILY OF CLASSIFICATIONS In nosology, a set of related classification systems describing different aspects of health problems. For example, the International Classification of Disease, the International Classification of Health Problems in Primary Care, the International Classification of Impairments, Disabilities and Handicaps, and the specialty subclassifications for oncology, psychiatry, etc. developed by WHO working groups constitute a "family of classifications."

FAMILY STUDY An epidemiologic study of a family or a group of families. The term has been used to describe surveillance of family groups, e.g., for tuberculosis. In genetics, investigation of families showing an unusual characteristic in order to determine whether the characteristic clusters in certain families and if so, why.

FAIRB, WILLIAM (1807-1883) A medical graduate who became the first compiler of abstracts (statistician) to the Registrar-General in the newly established General Register Office of England in 1839 and remained there for more than 40 years. In his *Annual Reports*, the combination of facts on death rates and vivid language drew attention to many inequalities of health and sickness experience between "healthy" and "unhealthy" districts in England. His many contributions to vital statistics and epidemiology are contained in his monograph *Vital Statistics* (London, 1885). These include a statement of the relationship between incidence and prevalence, the concepts of person-years, retrospective and prospective approaches, observed and expected numbers of events, the first workable nosology, and empirical laws about the natural history of epidemics.

FATALITY RATE The death rate observed in a designated series of persons affected by a simultaneous event, e.g., victims of a disaster. A term to be deprecated, because it can be confused with CASE FATALITY RATE.

FEASIBILITY STUDY Preliminary study to determine practicability of a proposed health program or procedure, or of a larger study, and to appraise the factors that may influence its practicability. See also PILOT STUDY.

fecundity The ability to produce live offspring. Fecundity is difficult to measure since it refers to the theoretical ability of a woman to conceive and carry a fetus to term. If a woman produces a live birth, it is known that she and her consort were fecund during some time in the past.

FERTILITY The actual production of live offspring. Stillbirths, fetal deaths, and abortions are not included in the measurement of fertility in a population. See also GRAVIDITY; PARITY.

FERTILITY RATE See GENERAL FERTILITY RATE.

FERTILITY RATIO A measure of the fertility of the population that restricts the denominator to the female population of appropriate age for childbearing. The fertility ratio is defined as

$$\text{Fertility ratio} = \frac{\text{Number of girls under 15 years of age}}{\text{Number of women in 15-49 age group}} \times 1000$$

(Not to be confused with GENERAL FERTILITY RATE.)

FETAL DEATH (Syn: stillbirth) Death prior to the complete expulsion or extraction from

its mother of a product of conception, irrespective of the duration of pregnancy. The death is indicated by the fact that after such separation the fetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles. Defined variously as death after the 20th or 28th week of gestation (the definition of the length of gestation varies between different jurisdictions, making this event difficult to compare internationally). See also LIVE BIRTH.

FETAL DEATH CERTIFICATE (Syn: certificate of stillbirth) A vital record registering a fetal death or stillbirth. Some health jurisdictions require the use of a fetal death certificate for all products of conception, whereas others require its use only in cases in which gestation has reached a particular duration, usually the 20th or the 28th week.

FETAL DEATH RATE (Syn: stillbirth rate) The number of fetal deaths in a year expressed as a proportion of the total number of births (live births plus fetal deaths) in the same year.

$$\text{Fetal death rate} = \frac{\text{Number of fetal deaths in a year}}{\text{Number of fetal deaths plus live births in the same year}} \times 1000$$

Note that the denominator is larger than for the FETAL DEATH RATIO and that the fetal death rate is therefore lower than the fetal death ratio, which is used in some jurisdictions. International comparisons of stillbirth or fetal death statistics will be flawed if the distinction is not appreciated.

FETAL DEATH RATIO A measure of fetal wastage, related to the number of live births. Defined as

$$\text{Fetal death ratio} = \frac{\text{Number of fetal deaths in a year}}{\text{Number of live births in the same year}}$$

(Can be expressed per 1000.)

FIELD SURVEY The planned collection of data in "the field," i.e., usually among noninstitutionalized persons in the general population. A method of establishing a relationship between two or more variables in a population in numerical terms by eliciting and collating information from existing sources (not only records but people who can say how they feel or what happened). See also CROSS-SECTIONAL STUDY.

FINLAY, CARLOS ALBERT (1833-1915) Cuban physician, initial investigator (1888-1891) of the role of *Aedes aegypti* (then known as *Culex fasciatus*) in the transmission of yellow fever. His experiments were unsatisfactory, but his theory was fully confirmed by the experiments of the team led by REED in which he took an active part.

FISHER'S EXACT TEST The test for association in a two-by-two table that is based upon the exact hypergeometric distribution of the frequencies within the table.

FISHING EXPEDITION Exploratory study to find clues and leads for further study. Although the term is sometimes used pejoratively, "fishing expeditions" may be done for worthwhile causes, e.g., to seek clues to the cause of a major life-threatening outbreak. A recent example was the initial investigation of Legionnaires' disease.

FITNESS This word has specific meanings in several fields related to epidemiology.

1. In population genetics, a measure of the relative survival and reproductive success of a given individual or phenotype, or population subgroup.
2. In health promotion, health risk appraisal, physical fitness is a set of attributes

that people have or achieve, that relate to their ability to perform physical activity. Intellectual and emotional fitness can also be described and to some extent measured.

FIXED COHORT A cohort in which membership is fixed by being present at some defining event ("zero time"); an example is the cohort comprising survivors of the atomic bomb exploded at Hiroshima. See also **CLOSED COHORT**.

FOLLOW-UP Observation over a period of time of an individual, group, or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables. See also **COHORT**.

FOLLOW-UP STUDY

1. A study in which individuals or populations, selected on the basis of whether they have been exposed to risk, received a specified preventive or therapeutic procedure, or possess a certain characteristic, are followed to assess the outcome of exposure, the procedure, or effect of the characteristic, e.g., occurrence of disease.

2. Synonym for **COHORT STUDY**.

FOMITES (singular, fomes) Articles that convey infection to others because they have been contaminated by pathogenic organisms. Examples include handkerchief, drinking glass, door handle, clothing, and toys.

FORCE OF MORBIDITY (Syn: hazard rate, instantaneous incidence density, instantaneous incidence rate, person-time incidence rate) Theoretical measure of the number of new cases that occur per unit of population-time, e.g., person-years at risk. This is a measure of the occurrence of disease at a point in time, t , defined mathematically as the limit, as Δt approaches zero, of

$$\frac{\text{Probability that a person well at time } t \text{ will develop the disease in the interval } t + \Delta t}{\Delta t}$$

The average value of this quantity over the interval t to $(t + \Delta t)$ can be estimated as

$$\frac{\text{Incident cases observed from } t \text{ to } (t + \Delta t)}{\text{Number of person-time units of experience observed from } t \text{ to } (t + \Delta t)}$$

FORCE OF MORTALITY (Syn: actuarial death rate) The hazard rate of the occurrence of death at a point in time t , i.e., the limit as Δt approaches zero, of the probability that an individual alive at time t will die by time $t + \Delta t$, divided by Δt . Distinct from cumulative death rate.

FORECASTING A method of estimating what may happen in the future that relies on extrapolation of existing trends (demographic, epidemiologic, etc.). It may be less useful than **SCENARIO BUILDING**, which has greater flexibility. For example, extrapolation of mortality trends for coronary heart disease in the early 1960s in the United States suggested that the mortality rates would continue to rise, perhaps indefinitely, whereas in fact the rates began to fall soon after that time.

FORTUITOUS RELATIONSHIP A relationship that occurs by chance and needs no further explanation.

FORWARD SURVIVAL ESTIMATE A procedure for estimating the age distribution at some later date by projecting forward an observed age distribution. The procedure uses survival ratios, often obtained from model life tables.

2503012414

FOURFOLD TABLE See **CONTINGENCY TABLE**.

FRACASTORIUS, GIROLAMO (1484–1553) Physician, poet, natural scientist, and a man of legends, said to have required surgery at birth to open fused lips and to have survived a lightning bolt that killed his mother while he was in her arms as an infant. He gave the word "syphilis" to the world in his mock-heroic poem, *Syphilis Sive Morbus Gallicus* (1530), which explicitly described the transmission of disease by acts of venery. In *De Contagione* (1546), he described transmission of infection by direct contact, by somites, and "at a distance," by which he meant droplets.

FRAMINGHAM STUDY Probably the best known cohort study of heart disease. Since 1949, samples of residents of Framingham, Massachusetts, have been subjects of investigations of risk factors in relation to the occurrence of heart disease and later, other outcomes.

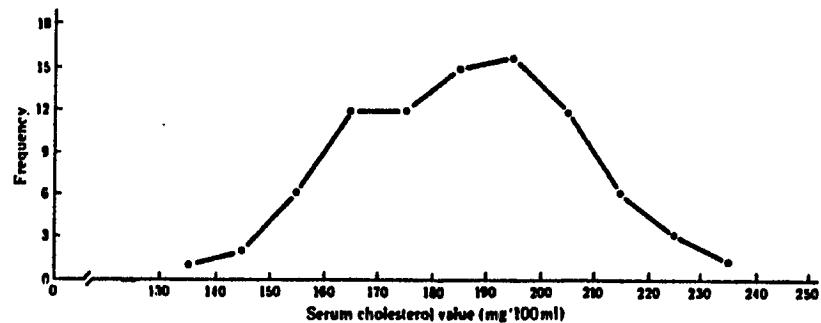
FRANK, JOHANN PIETER (1745–1821) Author of *System einer vollständigen medicinischen Polizei*, which established hygiene as a systematic science and contained many suggestions based on epidemiologic observations. In modern terminology, Frank was "Director-general of public health" to the Hapsburg empire in eighteenth century Vienna. His *System* contained many sensible rules for individual good health, and detailed specifications for public health practice.

FREQUENCY See **OCCURRENCE**.

FREQUENCY DISTRIBUTION See **DISTRIBUTION**.

FREQUENCY MATCHING See **MATCHING**.

FREQUENCY POLYGON A graphic illustration of a distribution, made by joining a set of points, for each of which the abscissa is the midpoint of the class and the ordinate, or height, is the frequency.



Frequency polygon. From Rimm et al., 1980.

FUNCTION A quality, trait, or fact that is so related to another as to be dependent upon and to vary with this other.

G

GALTON, FRANCIS (1822–1911) A founder of the modern science of human biology and the inventor of several statistical methods. Perhaps he is best known as the author of *Hereditary Genius* (1869), an analysis of physical and intellectual characteristics of successive generations of several hundred prominent families. Observing that offspring of parents of unusual talent, height, etc., tended toward average, he formulated the "Law of filial regression" (the origin of the term "regression"). His statistical approaches were refined and extended by his pupil, KARL PEARSON, the founder of modern BIOMETRY.

GAUSSIAN DISTRIBUTION See NORMAL DISTRIBUTION.

GAME THEORY A branch of mathematical logic concerned with the range of possible reactions to a particular strategy; each reaction can be assigned a probability and each reaction can lead to further action by the "adversary" in the game. Used mainly in systems analysis and such applications as war-gaming, game theory has occasional applications in disease surveillance and control. It is also one of the underlying theories used in clinical decision analysis.

GENE A sequence of DNA that codes for a particular protein product or that regulates other genes. Genes are the biological basis of heredity and occupy precisely defined locations on chromosomes.

GENE POOL The total of all genes possessed by reproductive members of a population.

GENERAL FERTILITY RATE A more refined measure of fertility than the crude birth rate. The denominator is restricted to the number of women of childbearing age (i.e., 15–44 or 15–49). Defined as

$$\text{General fertility rate} = \frac{\text{Number of live births in an area during a year}}{\text{Midyear female population age } 15\text{--}44 \text{ in same area in same year}} \times 1000$$

The upper age limit for this rate is 44 years in most jurisdictions.

GENERATION EFFECT (Syn: cohort effect) Variation in health status that arises from the different causal factors to which each birth cohort (see COHORT) in the population is exposed as the environment and society change. Each consecutive birth cohort is exposed to a unique environment that coincides with its life span.

GENERATION TIME The interval between receipt of infection by and maximal infectivity of the host. This applies to both clinical cases and inapparent infections.

With person-to-person transmission of infection, the interval between cases is determined by the generation time. See also INCUBATION PERIOD.

GENETIC DRIFT Random variation in gene frequency from generation to generation;

most often observed in small populations. The process of evolution through random statistical fluctuation of genetic composition of populations.

GENETIC EPIDEMIOLOGY The science that deals with the etiology, distribution and control of disease in groups of relatives, and with inherited causes of disease in populations.¹

¹Morton NE: *Outline of genetic epidemiology*. New York: Karger, 1982.

GENETIC LINKAGE Particular genes occupy specific sites in chromosomes, one member of each pair of chromosomes of course coming from each parent. When two genes are fairly close to each other in the same chromosome pair, they tend to be inherited together. Such genes are said to be linked, and the phenomenon is called genetic linkage.

GENETIC PENETRANCE The extent to which a genetically determined condition is expressed in an individual. This determines the frequency with which genetic effect is shown in a population.

GENETICS The branch of biology dealing with heredity and variation of individual members of a species. Its branches include population genetics, which overlaps epidemiology; therefore we include pertinent genetic terms in this dictionary.

GENOME The array of genes carried by an individual.

GEOGRAPHIC PATHOLOGY (Syn: medical geography) The comparative study of countries, or of regions within them, with regard to variations in morbidity/mortality. The (implied) aim of such study is usually to demonstrate that the variations are caused by or related to differences in the geographic environment.

GEOMETRIC MEAN See MEAN, GEOMETRIC.

GESTATIONAL AGE Strictly speaking, the gestational age of a fetus is the elapsed time since conception. However, as the moment when conception occurred is rarely known precisely, the duration of gestation is measured from the first day of the last normal menstrual period. Gestational age is expressed in completed days or completed weeks (e.g., events occurring 280–286 days after the onset of the last normal menstrual period are considered to have occurred at 40 weeks of gestation).

Measurements of fetal growth, as they represent continuous variables, are expressed in relation to a specific week of gestational age (e.g., the mean birth weight for 40 weeks is that obtained at 280–286 days of gestation on a weight-for-gestational-age curve). Some specified variations of gestational age are: *Prem*: Less than 37 completed weeks (less than 259 days). *Term*: From 37 to less than 42 completed weeks (259–293 days). *Postterm*: Forty-two completed weeks or more (294 days or more).

"GOLD STANDARD" A jargon term, used to describe a method, procedure, or measurement that is widely accepted as being the best available. Often used to compare with new methods.

GOLDBERGER, JOSEPH (1874–1927) A U.S. Public Health Service physician. Responsible for a brilliant series of investigations of pellagra. After logical deductions led him to reject the prevailing view that pellagra had an infectious origin, he conducted studies in several rural communities and in institutions, leading conclusively to the demonstration that pellagra was a dietary deficiency disease.

GOMPERTZ'S LAW The proportionate relationship of mortality to age. Mortality is high during the first year of life (infancy), drops to its lowest level in childhood, and gradually climbs during the third and fourth decade. After age 35 or 40, the increase in mortality with age tends to be logarithmic for the remainder of the life span, i.e., the relative increase in mortality in each successive age class (of equal

size) is about constant. This law was first enunciated by the demographer Benjamin Gompertz, on the basis of survival curves in English villages in the 1840s.

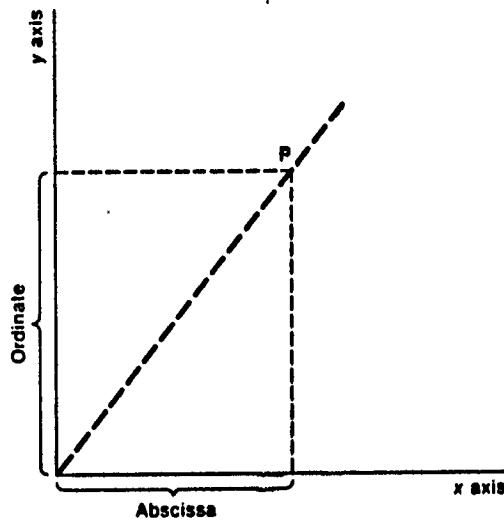
GONADOTROPHIC CYCLE One complete round of ovarian development in the mosquito (or other insect vector) from the time when the blood meal is taken to the time when the fully developed eggs are laid.

GOODNESS OF FIT Degree of agreement between an empirically observed distribution and a mathematical or theoretical distribution.

GOODNESS OF FIT TEST A statistical test of the hypothesis that data have been randomly sampled or generated from a population that follows a particular theoretical distribution or model. The most common such tests are chi-square tests.

GRADIENT OF INFECTION The variety of host responses to infection ranging from inapparent infection to fatal illness.

GRAPH Visual display of the relationship between variables; the values of one set of variables are plotted along the horizontal or *x* axis, of a second variable, along the vertical or *y* axis. Three-dimensional graphs of relationships between three variables can be represented and comprehended visually in two dimensions. The relationship between *x* and *y* may be linear, exponential, logarithmic, etc. See also **AXIS**, **ABSCISSA**, **ORDINATE**. "Graph" is also a descriptive term for histograms, bar charts, etc.



Graph showing abscissa, ordinate, and locus of a point, *P*, in relation to *x* and *y* axis.

GRAUNT, JOHN (1620–1674) By profession a haberdasher, he was a member of the small community of scholars and natural scientists in London who were Fellows of the Royal Society in its early years and who made important contributions to the natural sciences. Graunt studied the **BILLS OF MORTALITY** and used them to conduct the first analytic studies of vital statistics, identifying differences in mortality rates between the sexes, between city and country folk, and recording all in *Natural and political observations mentioned in a following index and made upon the Bills of Mortality* (London, 1662).

GRAVIDITY The number of pregnancies (completed or incomplete) experienced by a woman.

GREENWOOD, MAJOR (1888–1949) Medical epidemiologist, trained in statistics by Karl Pearson; Greenwood was the first professor of epidemiology at the London School of Hygiene and Tropical Medicine. He inspired a whole generation of British epidemiologists, introducing to the subject a level of mathematical reasoning and statistical rigor it had not previously known. Author of many papers and several monographs, best known of which is *Epidemics and Crowd Diseases* (London, 1933).

GROSS REPRODUCTION RATE The average number of female children a woman would have if she survived to the end of her childbearing years and if, throughout that period, she were subject to a given set of age-specific fertility rates and a given sex ratio at birth. This rate provides a measure of replacement fertility in the absence of mortality. See also **NET REPRODUCTION RATE**.

GROWTH RATE OF POPULATION A measure of population growth (in the absence of migration) comprising addition of newborns to the population and subtraction of deaths. The result, known as *natural rate of increase*, is calculated as

$$\frac{\text{Live births during the year} - \text{deaths during the year}}{\text{Midyear population}} \times 100$$

Alternatively, it is the difference between crude birth rate and crude death rate.

H

HACKETT SPLEEN CLASSIFICATION A numerical means of recording the size of an enlarged spleen, especially in malaria. This is a 6-point scale of 0 (no enlargement) to 5 (enlarged to umbilicus or larger). See *Terminology of Malaria and of Malaria Eradication*. Geneva: WHO, 1963, pp. 40-41.

HALO EFFECT

1. The effect (usually beneficial) that the manner, attention, and caring of a provider have on a patient during a medical encounter regardless of what medical procedures or services the encounter involves. See also PLACEBO, PLACEBO EFFECT.
2. The influence upon an observation of the observer's perception of the characteristics of the individual observed (other than the characteristic under study) or the influence of the observer's recollection or knowledge of findings on a previous occasion.

HANDICAP Reduction in a person's capacity to fulfill a social role as a consequence of an IMPAIRMENT, inadequate training for the role, or other circumstances. Applied to children, the term usually refers to the presence of an impairment or other circumstance that is likely to interfere with normal growth and development or with the capacity to learn. See also INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES, AND HANDICAPS for the official WHO definition.

HAPHAZARD SAMPLE Selection of a group of persons for study without thought as to whether they are representative of the population. The word "haphazard" here implies selection based on a mixture of criteria such as convenience, accessibility, turning up at the time an investigation or study is in progress, and belonging to some existing list or registry, etc. Because they have an unknown chance of being unrepresentative of the population, haphazard samples are unsatisfactory for generalization.

HARDY-WEINBERG LAW The principle that both gene and genotype frequencies will remain in equilibrium in an infinitely large population in the absence of mutation, migration, selection, and nonrandom mating. If p is the frequency of one allele and q is the frequency of another and $p+q=1$, then p^2 is the frequency of homozygotes for the allele, q^2 is the frequency of homozygotes for the other allele, and $2pq$ is the frequency of heterozygotes.

HARMONIC MEAN See MEAN, HARMONIC.

HAWTHORNE EFFECT The effect (usually positive or beneficial) of being under study upon the persons being studied; their knowledge of the study often influences their behavior. The name derives from work studies by Whitehead, Dickson, Roethlisberger, and others, in the Western Electric Plant, Hawthorne, Illinois, reported by Elton Mayo in *The Social Problems of an Industrial Civilization* (London: Routledge, 1949).

2503012417

HAZARD A factor or exposure that may adversely affect health.

HAZARD RATE (Syn: force of morbidity, instantaneous incidence rate) A theoretical measure of the risk of occurrence of an event, e.g., death, new disease, at a point in time, t , defined mathematically as the limit, as Δt approaches zero, of the probability that an individual well at time t will experience the event by $t+\Delta t$, divided by Δt .

HEALTH The World Health Organization (WHO) described *health* in the preamble to its constitution as, "A state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity." The WHO description of health has been criticized because of the difficulty of defining and measuring "complete" well-being.

There are several other definitions, including the following:

A state of dynamic balance in which an individual's or a group's capacity to cope with all the circumstances of living is at an optimum level.

A state characterized by anatomical, physiological and psychological integrity, ability to perform personally valued family, work and community roles; ability to deal with physical, biological, psychological and social stress; a feeling of well-being; and freedom from the risk of disease and untimely death.

Rene Dubus offered the following definition: "A modus vivendi enabling imperfect men to achieve a rewarding and not too painful existence while they cope with an imperfect world."

The word "health" is derived from the Old English *Hal*, meaning hale, whole, sound in wind and limb.

HEALTH BEHAVIOR The combination of knowledge, practices, and attitudes that together contribute to motivate the actions we take regarding health. Health behavior may promote and preserve good health, or if the behavior is harmful, e.g., tobacco smoking, may be a determinant of disease. This combination of knowledge, practices, and attitudes has been described and discussed by several writers, notably Becker.¹ See also *Illness behavior*.

¹ Becker MH (ed): *The Health Belief Model and Personal Health Behavior*. Thorofare NJ: Slack, 1974.

HEALTH CARE Those services provided to individuals or communities by agents of the health services or professions, for the purpose of promoting, maintaining, monitoring, or restoring health. Health care is broader than, and not limited to medical care, which implies therapeutic action by or under the supervision of a physician. The term is sometimes extended to include self-care.

HEALTH EDUCATION The process by which individuals and groups of people learn to behave in a manner conducive to the promotion, maintenance, or restoration of health.

HEALTH INDEX A numerical indication of the health of a given population derived from a specified composite formula. The components of the formula may be INFANT MORTALITY RATES, INCIDENCE RATES for particular disease, or other HEALTH INDICATORS.

HEALTH INDICATOR A variable, susceptible to direct measurement, that reflects the state of health of persons in a community. Examples include infant mortality rates, incidence rates based on notified cases of disease, disability days, etc. These measures may be used as components in the calculation of a HEALTH INDEX.

HEALTH PROMOTION The process of enabling people to increase control over and improve their health. It involves the population as a whole in the context of their everyday lives, rather than focusing on people at risk for specific diseases, and is directed toward action on the determinants or causes of health.

HEALTH RISK APPRAISAL (HRA) (Syn: health hazard appraisal [HHA]) A generic term applied to methods for describing an individual's chances of becoming ill or dying from selected causes. The many versions now available share several common features: Starting from the average risk of death for the individual's age and sex, a consideration of various lifestyle and physical factors indicates whether the individual is at greater or less than average risk of death from the commonest causes of death for his age and sex. All methods also indicate what reduction in risk could be achieved by altering any of the causal factors (such as cigarette smoking) that the individual could modify.

The premise underlying such methods is that information on the extent to which an individual's characteristics, habits, and health practices are influencing his future risk of dying will assist health care workers in counseling their patients.

HEALTH SERVICES Services that are performed by health care professionals, or by others under their direction, for the purpose of promoting, maintaining, or restoring health. In addition to personal health care, health services include measures for health protection and health education.

HEALTH SERVICES RESEARCH The integration of epidemiologic, sociological, economic, and other analytic sciences in the study of health services. Health services research is usually concerned with relationships between NEED, DEMAND, supply, use, and OUTCOME of health services. The aim of health services research is evaluation; several components of evaluative health services research are distinguished, viz:

Evaluation of *structure*, concerned with resources, facilities, and manpower.

Evaluation of *process*, concerned with matters such as where, by whom, and how health care is provided.

Evaluation of *output*, concerned with the amount and nature of health services provided.

Evaluation of *outcome*, concerned with the results, i.e., whether persons using health services experience measurable benefits such as improved survival or reduced disability.

HEALTH STATISTICS Aggregated data describing and enumerating attributes, events, behaviors, services, resources, outcomes, or costs related to health, disease, and health services. The data may be derived from survey instruments, medical records, and administrative documents. VITAL STATISTICS are a subset of health statistics.

HEALTH STATUS INDEX A set of measurements designed to detect short-term fluctuations in the health of members of a population; these measurements generally include physical function, emotional well-being, activities of daily living, feelings, etc. Most indexes require the use of carefully composed questions designed with reference to matters of fact rather than shades of opinion. The results are usually expressed by a numerical score that gives a profile of the well-being of the individual.

HEALTH SURVEY A survey designed to provide information on the health status of a population. It may be descriptive, exploratory, or explanatory. See also MORBIDITY SURVEY.

HEALTHY WORKER EFFECT A phenomenon observed initially in studies of occupational diseases: Workers usually exhibit lower overall death rates than the general population, due to the fact that the severely ill and disabled are ordinarily excluded from employment. Death rates in the general population may be inappropriate for comparison if this effect is not taken into account.

HEBDOMADAL MORTALITY RATE The mortality rate in the first week of life; the denominator is the number of live births in a year.

HENLE-KOCH POSTULATES See KOCH'S POSTULATES.

HERD IMMUNITY The immunity of a group or community. The resistance of a group to

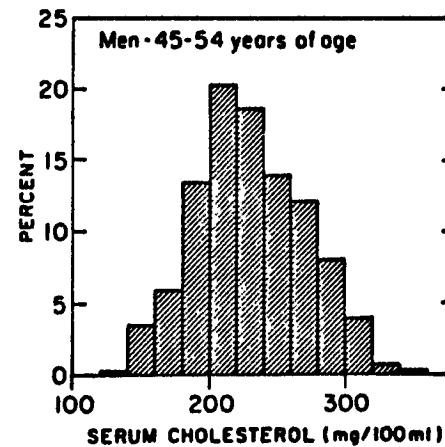
invasion and spread of an infectious agent, based on the resistance to infection of a high proportion of individual members of the group. The resistance is a product of the number susceptible and the probability that those who are susceptible will come into contact with an infected person. In the herd immunity equation, "probability of contact" is the intervening factor that reduces susceptibility to infection among group members to less than that anticipated from their susceptibility as unrelated individuals.

HETEROSEDASTICITY Nonconstancy of the variance of a measure over the levels of the factors under study.

HIBERNATION See VECTOR-BORNE INFECTION.

HIPPOCRATES OF COS (c 460-370 BC) Greek physician, "Father of Medicine," responsible for careful clinical observation of many important and common diseases—tetanus, mumps, puerperal septicemia, etc. His writings contain important epidemiologic observations, as in the books *Airs, Waters, Places*, and *Epidemics*. His *Aphorisms* also demonstrate considerable empirical epidemiologic knowledge.

HISTOGRAM A graphic representation of the frequency distribution of a variable. Rectangles are drawn in such a way that their bases lie on a linear scale representing different intervals, and their heights are proportional to the frequencies of the values within each of the intervals. See also BAR DIAGRAM.



Histogram. From National Center for Health Statistics, 1978.

HISTORICAL COHORT STUDY (Syn: historical prospective study, nonconcurrent prospective study, prospective study in retrospect) A COHORT STUDY conducted by reconstructing data about persons at a time or times in the past. This method uses existing records about the health or other relevant aspects of a population as it was at some time in the past and determines the current (or subsequent) status of members of this population with respect to the condition of interest. Different levels of past exposure to risk factor(s) of interest must be identifiable for subsets of the population. See also COHORT STUDY.

HISTORICAL CONTROL Control subject(s) for whom data were collected at a time preceding that at which the data are gathered on the group being studied. Because of differences in exposures etc., use of historical controls can lead to bias in analysis.

HOGBEN NUMBER A unique personal identifying number constructed by using a sequence of digits for birthdate, sex, birthplace, and other identifiers. Suggested by the English mathematician Lancelot Hogben. Used in primary care epidemiology in some countries and usable in RECORD LINKAGE. See also IDENTIFICATION NUMBER; SOUNDEX CODE.

HOLMES, OLIVER WENDELL (1809–1894) Physician, poet, philosopher, autocrat ("of the Breakfast Table"), and crusader against puerperal fever. He argued that this was conveyed to patients by the contaminated hands and clothes of attending physicians and recommended washing the hands and changing clothes as a way to prevent it. Unlike SEMMELWEIS, he succeeded in convincing the medical profession. His correct belief was recorded in a paper, "The Contagiousness of Puerperal Fever."¹

¹ *N Eng Q J Med Surg* 1:503–530, 1842–43.

HOLOENDEMIC DISEASE A disease for which a high prevalent level of infection begins early in life and affects most of the child population, leading to a state of equilibrium such that the adult population shows evidence of the disease much less commonly than do the children. Malaria in many communities is a holendemic disease.

HOLOMANTIC INFECTION See common source epidemic.

HOMOSCEDASTICITY Constancy of the variance of a measure over the levels of the factors under study.

HOSPITAL-ACQUIRED INFECTION See NOSOCOMIAL INFECTION.

HOSPITAL DISCHARGE ABSTRACT SYSTEM Abstraction of MINIMUM DATA SET from hospital charts for the purpose of producing summary statistics about hospitalized patients. Examples include the Hospital Inpatient Enquiry (HIPE) and Professional Activity Study (PAS). The statistical tabulations commonly include length of stay by final diagnosis, surgical operations, specified hospital service (i.e., medical, surgical, gynecological, etc.) and also give outcomes such as "death" and "discharged alive from hospital." This system cannot generally be used for epidemiologic purposes as it is not possible to infer representativeness or to generalize; this is because the data usually lack a defined denominator and the same person may be counted more than once in the event of two or more HOSPITAL SEPARATIONS in the period of study.

HOSPITAL INPATIENT ENQUIRY (HIPE) Statistical tables of a 10% sample of hospital patients in England and Wales, showing class of hospital, diagnosis, length of stay, outcomes, etc.

HOSPITAL SEPARATION A term used in commentaries on hospital statistics to describe the departure of a patient from hospital without distinguishing whether the patient departed alive or dead (the distinction is unimportant so far as the statistics of hospital activity such as bed occupancy are concerned).

HOST

1. A person or other living animal, including birds and arthropods, that affords subsistence or lodgment to an infectious agent under natural conditions. Some protozoa and helminths pass successive stages in alternate hosts of different species. Hosts in which the parasite attains maturity or passes its sexual stage are primary or definitive hosts; those in which the parasite is in a larval or asexual state are secondary or intermediate hosts. A transport host is a carrier in which the organism remains alive but does not undergo development.¹
2. In an epidemiologic context, the host may be the population or group; biological, social, and behavioral characteristics of this group that are relevant to health are called "host factors."

¹ Benenson, *op. cit.*

HOUSEHOLD One or more persons who occupy a dwelling, i.e., a place that provides

shelter, cooking, washing, and sleeping facilities; may or may not be a family. The term is also used to describe the dwelling unit in which the persons live.

HOUSEHOLD SAMPLE SURVEY A survey of persons in a sample of households. This, in many variations, is a favored method of gathering data for health-related and for many other purposes. The households may be sampled in any of several ways, e.g., by cluster, use of random numbers in relation to numbered dwelling units. The survey may be conducted by interview, telephone survey, or self-completed responses to present questions. The method is used in developing nations as well as in the industrial world.

HUMAN BLOOD INDEX Proportion of insect vectors found to contain human blood.

HUMAN ECOLOGY See ECOLOGY.

HUMAN IMMUNODEFICIENCY VIRUS (HIV) The pathogenic organism responsible for the acquired immunodeficiency syndrome (AIDS); formerly or also known as the lymphadenopathy virus (LAV), the name given by the original French discoverers Montagnier et al.¹ in 1983, or the human T-cell lymphotropic virus, type III (HTLV-III), the name given by Gallo et al.² to the virus they reported in 1984.

¹ Barre-Sinoussi F, Chermann JC, Rey F, et al.: Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220:868–871, 1983.

² Gallo RC, Salahuddin SZ, Popovic M, et al.: Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* 224:500–503, 1984.

HYPERENDEMIC DISEASE A disease that is constantly present at a high incidence and/or prevalence rate and affects all age groups equally.

HYPERGEOMETRIC DISTRIBUTION The exact probability distribution of the frequencies in a two-by-two contingency table, conditional on the marginal frequencies being fixed at their observed levels.

HYPOTHESIS

1. A supposition, arrived at from observation or reflection, that leads to refutable predictions.
2. Any conjecture cast in a form that will allow it to be tested and refuted.

See also NULL HYPOTHESIS.

I

IATROGENIC DISEASE Illness resulting from a physician's professional activity, or from the professional activity of other health professionals.

ICD See INTERNATIONAL CLASSIFICATION OF DISEASE.

ICEBERG PHENOMENON That portion of disease that remains unrecorded or undetected despite physicians' diagnostic endeavors and community disease surveillance procedures is referred to as the "submerged portion of the iceberg." Detected or diagnosed disease is the "tip of the iceberg." The submerged portion comprises disease not medically attended, medically attended but not accurately diagnosed, and diagnosed but not reported.¹

¹Last JM: The iceberg. *Lancet*, 2:28-31, 1963.

ICHPPC See INTERNATIONAL CLASSIFICATION OF HEALTH PROBLEMS IN PRIMARY CARE.

IDENTIFICATION NUMBER, IDENTIFYING NUMBER Unique number given to every individual at birth or at some other milestone. Sweden has a system based on a sequence of digits for birthdate, sex, birthplace, and additional digits for each individual. Other systems, e.g., National Insurance number in the United Kingdom, Social Security number in the United States, and Social Insurance number in Canada, are sometimes used but are neither universal nor unique, being sometimes applied to whole families or at least to more than one individual. See also HOGBEN NUMBER; SOUNDEX CODE.

IDIOSYNCRASY Webster's Dictionary defines this as a distinctive characteristic or peculiarity of an individual. In pharmacoepidemiology, it means an abnormal reaction, sometimes genetically determined, following the administration of a medication.

ILLNESS See DISEASE.

ILLNESS BEHAVIOR Conduct of persons in response to abnormal body signals. Such behavior influences the manner in which a person monitors his body, defines and interprets his symptoms, takes remedial actions, and uses the health care system. See also HEALTH BEHAVIOR.

IMMUNITY, ACQUIRED Resistance acquired by a host as a result of previous exposure to a natural PATHOGEN or foreign substance for the host, e.g., immunity to measles resulting from a prior infection with measles virus.

IMMUNITY, ACTIVE Resistance developed in response to stimulus by an antigen (infecting agent or vaccine) and usually characterized by the presence of antibody produced by the host.

IMMUNITY, NATURAL Species-determined inherent resistance to a disease agent, e.g., resistance of man to virus of canine distemper.

IMMUNITY, PASSIVE Immunity conferred by an antibody produced in another host and acquired naturally by an infant from its mother or artificially by administration of an antibody-containing preparation (antisera or immune globulin).

IMMUNITY, SPECIFIC A state of altered responsiveness to a specific substance acquired

through immunization or natural infection. For certain diseases (e.g., measles, chickenpox) this protection generally lasts for the life of the individual.

IMMUNIZATION (Syn: vaccination) Protection of susceptible individuals from communicable disease by administration of a living modified agent (as in yellow fever), a suspension of killed organisms (as in whooping cough), or an inactivated toxin (as in tetanus). Temporary passive immunization can be produced by administration of antibody in the form of immune globulin in some conditions.

IMPAIRMENT A physical or mental defect at the level of a body system or organ. See also INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES, AND HANDICAPS for the official WHO definition.

INAPPARENT INFECTION (Syn: subclinical infection) The presence of infection in a host without occurrence of recognizable clinical signs or symptoms. Of epidemiologic significance because hosts so infected, though apparently well, may serve as silent or inapparent disseminators of the infectious agent. See also DISEASE, PRECLINICAL; DISEASE, SUBCLINICAL; VECTOR-BORNE INFECTION.

INCEPTION RATE The rate at which new spells of illness occur in a population; a term applied principally to short-term spells of illness such as acute respiratory infections, and preferred by some epidemiologists because an annual incidence rate for such conditions may exceed the numbers in the population at risk.

INCIDENCE (Syn: incident number) The number of instances of illness commencing, or of persons falling ill, during a given period in a specified population.¹ More generally, the number of new events, e.g., new cases of a disease in a defined population, within a specified period of time. The term incidence is sometimes used to denote INCIDENCE RATE.

¹Prevalence and Incidence. *WHO Bull* 35:783-784, 1966.

INCIDENCE DENSITY The person-time incidence rate; sometimes used to describe the hazard rate. See FORCE OF MORBIDITY.

INCIDENCE-DENSITY RATIO (IDR) The ratio of two incidence densities. See also RATE RATIO.

INCIDENCE RATE The rate at which new events occur in a population. The numerator is the number of new events that occur in a defined period; the denominator is the population at risk of experiencing the event during this period, sometimes expressed as person-time. The incidence rate most often used in public health practice is calculated by the formula

$$\frac{\text{Number of new events in specified period}}{\text{Number of persons exposed to risk during this period}} \times 10^4$$

In a dynamic population, the denominator is the average size of the population, often the estimated population at the mid-period. If the period is a year, this is the annual incidence rate. This rate is an estimate of the person-time incidence rate, i.e., the rate per 10⁴ person-years. If the rate is low, as with many chronic diseases, it is also a good estimate of the cumulative incidence rate. In follow-up studies with no censoring, the incidence rate is calculated by dividing the number of new cases in a specified period by the initial size of the cohort of persons being followed; this is equivalent to the cumulative incidence rate during the period. If the number of new cases during a specified period is divided by the sum of the person-time units at risk for all persons during the period, the result is the person-time incidence rate.

INCIDENCE STUDY See COHORT STUDY.

INCIDENT NUMBER See INCIDENCE.

INCUBATION PERIOD

1. The time interval between invasion by an infectious agent and appearance of the first sign or symptom of the disease in question.
2. In a VECTOR, the period between entry of the infectious agent into the vector and the time at which the vector becomes infective; i.e., transmission of the infectious agent from the vector to a fresh final host is possible (extrinsic incubation period).

INDEPENDENCE Two events are said to be independent if the occurrence of one is in no way predictable from the occurrence of the other. Two variables are said to be independent if the distribution of values of one is the same for all values of the other. Independence is the antonym of ASSOCIATION.

INDEPENDENT VARIABLE

1. The characteristic being observed or measured that is hypothesized to influence an event or manifestation (the dependent variable) within the defined area of relationships under study; that is, the independent variable is not influenced by the event or manifestation but may cause it or contribute to its variation.
2. In statistics, an independent variable is one of (perhaps) several variables that appear as arguments in a regression equation.

INDEX In epidemiology and related sciences, this word usually means a rating scale, e.g., a set of numbers derived from a series of observations of specified variables. Examples include the many varieties of health status index, scoring systems for severity or stage of cancer, heart murmurs, mental retardation, etc.

INDEX CASE The first case in a family or other defined group to come to the attention of the investigator. See also PROPOSITUS.

INDEX GROUP (Syn: index series)

1. In an experiment, the group receiving the experimental regimen.
2. In a case control study, the cases.
3. In a cohort study, the exposed group.

INDICATOR VARIABLE In statistics, a variable taking only one of two possible values, one (usually 1) indicating the presence of a condition, and the other (usually zero) indicating absence of the condition. Used mainly in REGRESSION ANALYSIS.

INDIRECT ADJUSTMENT See standardization.

INDIVIDUAL VARIATION Two types are distinguished:

1. *Intraindividual variation*: The variation of biological variables within the same individual, depending upon circumstances such as the phase of certain body rhythms and the presence or absence of emotional stress. These variables do not have a precise value, but rather a range. Examples include diurnal variation in body temperature, fluctuation of blood pressure, blood sugar, etc.
2. *Interindividual variation*: As used by Darwin, the term means variation between individuals. This is the preferred usage; the first usage is better described as personal variation.

INDUCTION PERIOD The period required for a specific cause to produce disease. More precisely, the interval from the causal action of a factor to the initiation of the disease. For example, a span of many years may pass between (presumably) radiation-induced mutations and the appearance of leukemia; this span would be the induction period for radiogenic leukemia. See also INCUBATION PERIOD; LATENT PERIOD.

2503012421

INDUSTRIAL HYGIENE The science and art devoted to recognition, evaluation, and control of those environmental factors or stresses arising from or in the workplace, which may cause sickness, impaired health, and well-being, or significant discomfort and inefficiency among workers or among persons in the community. Alternatively, the profession that anticipates and controls unhealthy conditions of work to prevent illness among employees.

INFANT MORTALITY RATE (IMR) A measure of the yearly rate of deaths in children less than one year old. The denominator is the number of live births in the same year. Defined as

$$\text{Infant mortality rate} = \frac{\text{Number of deaths in a year of children less than 1 year of age}}{\text{Number of live births in the same year}} \times 1000$$

This is often quoted as a useful indicator of the level of health in a community.

INFECTIBILITY The host characteristic or state in which the host is capable of being infected. See also INFECTIOUSNESS; INFECTIVITY.

INFECTION (Syn: colonization) The entry and development or multiplication of an infectious agent in the body of man or animals. Infection is not synonymous with infectious disease; the result may be inapparent or manifest. The presence of living infectious agents on exterior surfaces of the body is called "infestation" (e.g., pediculosis, scabies). The presence of living infectious agents upon articles of apparel or soiled articles is not infection, but represents CONTAMINATION of such articles. See also INAPPARENT INFECTION; TRANSMISSION OF INFECTION.

INFECTION, GRADIENT OF The range of manifestations of illness in the host reflecting the response to an infectious agent, which extends from death at one extreme to inapparent infection at the other. The frequency of these manifestations varies with the specific infectious disease. For example, human infection with the virus of rabies is almost invariably fatal, whereas a high proportion of persons infected in childhood with the virus of hepatitis A, experience a subclinical or mild clinical infection.

INFECTION, LATENT PERIOD OF The time between initiation of infection and first shedding or excretion of the agent.

INFECTION, SUBCLINICAL See INAPPARENT INFECTION.

INFECTIOUS DISEASE See COMMUNICABLE DISEASE.

INFECTIOUSNESS A characteristic of the disease that concerns the relative ease with which it is transmitted to other hosts. A droplet spread disease, for instance, is more infectious than one spread by direct contact. The characteristics of the portals of exit and entry are thus also determinants of infectiousness, as are the agent characteristics of ability to survive away from the host, and of infectivity.

INFECTIVITY

1. The characteristic of the disease agent that embodies capability to enter, survive, and multiply in the host. A measure of infectivity is the secondary attack rate.
2. The proportion of exposures, in defined circumstances, that results in infection.

INFERENCE The process of passing from observations and axioms to generalizations. In statistics, the development of generalization from sample data, usually with calculated degrees of uncertainty.

INFESTATION The development on (rather than in) the body of a pathogenic agent, e.g., body lice. Some authors use the term also to describe invasion of the gut by parasitic worms.

INFORMATION SYSTEM A combination of vital and health statistical data from multiple sources, used to derive information about the health needs, health resources, costs, use of health services, and outcomes of use by the population of a specified jurisdiction. The term may also describe the automatic release from computers of stored information in response to programmed stimuli. For example, parents can be notified when their children are due to receive booster doses of an immunizing agent against infectious disease.

INFORMED CONSENT Voluntary consent given by a subject or by a person responsible for a subject (e.g., a parent) for participation in an investigation, immunization program, treatment regimen, etc., after being informed of the purpose, methods, procedures, benefits, and risks. Awareness of risk is necessary for any subject to make an informed choice. The term also refers to consent for medical care.

INOCULATION See VACCINATION.

INPUT

1. The sum total of resources and energies purposefully engaged in order to intervene in the spontaneous operation of a system.
2. The basic resources required in terms of manpower, money, materials, and time.

INSTANTANEOUS INCIDENCE RATE See FORCE OF MORBIDITY.

INSTRUMENTAL ERROR Error due to faults arising in any or in all aspects of a measuring instrument, i.e., calibration, accuracy, precision, etc. Also applied to error arising from impure reagents, wrong dilutions, etc.

INTERACTION

1. The interdependent operation of two or more causes to produce or prevent an effect. *Biological interaction* means the interdependent operation of two or more causes to produce, prevent, or control disease. See also ANTAGONISM; SYNERGISM.
2. Differences in the effects of one or more factors according to the level of the remaining factor(s). See also EFFECT MODIFIER.
3. In statistics, the necessity for a product term in a linear model.

INTERMEDIATE VARIABLE (Syn: contingent variable, intervening [causal] variable, mediator variable) A variable that occurs in a causal pathway from an independent to a dependent variable. It causes variation in the dependent variable, and itself is caused to vary by the independent variable. Such a variable is statistically associated with both the independent and dependent variables.

INTERNAL VALIDITY See VALIDITY, STUDY.

INTERNATIONAL CLASSIFICATION OF DISEASE (ICD) The classification of specific conditions and groups of conditions determined by an internationally representative group of experts who advise the World Health Organization, which publishes the complete list in a periodically revised book, the (*Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death*). Every disease entity is assigned a number. There are 17 major divisions (chapters) and a hierarchical arrangement of subdivisions (rubrics) within each. Some chapters are "etiological," e.g., Infective and Parasitic Conditions; more relate to body systems, e.g., Circulatory System; and some to classes of condition, e.g., neoplasms, injury (violence). The heterogeneity of categories reflects prevailing uncertainties about causes of disease (and classifi-

cation in relation to causes). The Ninth Revision of the Manual (ICD-9) was published by WHO in 1977, after ratification in 1976.

INTERNATIONAL CLASSIFICATION OF HEALTH PROBLEMS IN PRIMARY CARE (ICHPPC) A classification of diseases, conditions, and other reasons for attendance for primary care. May be used for labeling conditions in problem-oriented records as used by primary care health workers. This classification is an adaptation of the ICD but makes more allowance for the diagnostic uncertainty that prevails in primary care. This classification is now in its second revision (ICHPPC-2). See also PROBLEM-ORIENTED MEDICAL RECORD.

INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES, AND HANDICAPS (ICIDH) First published by WHO in 1980, this is an attempt to produce a systematic taxonomy of the consequences of injury and disease.

An *impairment* is defined in ICIDH as any loss or abnormality of psychological, physiological, or anatomical structure or function. It is concerned with abnormalities of body structure and appearance and with organ or system function resulting from any cause; in principle, impairments represent disturbances at the organ level.

A *disability* is defined in ICIDH as any restriction or lack (resulting from an impairment) of ability to perform an activity in a manner or within the range considered normal for a human being. The term disability reflects the consequences of impairment in terms of functional performance and activity by the individual; disabilities thus represent disturbances at the level of the person.

A *handicap* is defined in ICIDH as a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfillment of a role that is normal (depending on age, sex, and social and cultural practice) for that individual. The term handicap thus reflects interaction with and adaptation to the individual's surroundings.

INTERNATIONAL COMPARISON See GEOGRAPHIC PATHOLOGY. See also CROSS-CULTURAL STUDY.

INTERNAL VALIDITY See VALIDITY, STUDY.

INTERPOLATE, INTERPOLATION To predict the value of variates within the range of observations; the resulting prediction.

INTERVAL INCIDENCE DENSITY See PERSON-TIME INCIDENCE RATE.

INTERVAL SCALE See MEASUREMENT SCALE.

INTERVENING CAUSE See INTERMEDIATE VARIABLE.

INTERVENING VARIABLE

1. Synonym for INTERMEDIATE VARIABLE.
2. A variable whose value is altered in order to block or alter the effect(s) of another factor.

See also CAUSALITY, FACTORS IN.

INTERVENTION STUDY An epidemiologic investigation designed to test a hypothesized cause-effect relationship by modifying a supposed causal factor in a population.

INTERVIEW SCHEDULE The precisely designed set of questions used in an interview. See also SURVEY INSTRUMENT.

INVOLUNTARY SMOKING (Syn: passive smoking) The inhalation by nonsmokers of tobacco smoke left in the air by smokers; includes both smoke exhaled by smokers and smoke released directly from burning tobacco into ambient air; the latter is called sidestream smoke and contains higher proportions of toxic and other carcinogenic substances than exhaled smoke. The adjective "involuntary" is preferable to "passive" as the latter implies acquiescence—increasingly, nonsmokers are anything

but acquiescent about this form of air pollution. "Passive" is, however, customary WHO usage.

ISLAND POPULATION A group of individuals isolated from larger groups and possessing a relatively limited gene pool; alternatively, a group that is immunologically isolated and may therefore be unduly susceptible to infection with alien pathogens.

ISOLATE (noun) Term used in genetics to describe a subpopulation (generally small) in which matings take place exclusively with other members of the same subpopulation.

ISOLATION

1. In microbiology, the separation of an organism from others, usually by making serial cultures.
2. Separation, for the period of communicability, of infected persons or animals from others in such places and under such conditions as to prevent or limit the direct or indirect transmission of the infectious agent from those infected to those who are susceptible or who may spread the agent to others. *Control of Communicable Disease in Man*¹ lists seven categories of isolation as follows:
 - a. **Strict isolation:** This category is designed to prevent transmission of highly contagious or virulent infections that may be spread by both air and contact. The specifications, in addition to those above, include a private room and the use of masks, gowns, and gloves for all persons entering the room. Special ventilation requirements with the room at negative pressure to surrounding areas are desirable.
 - b. **Contact isolation:** For less highly transmissible or serious infections, for diseases or conditions that are spread primarily by close or direct contact. In addition to the basic requirements, a private room is indicated but patients infected with the same pathogen may share a room. Masks are indicated for those who come close to the patient, gowns are indicated if soiling is likely, and gloves are indicated for touching infectious material.
 - c. **Respiratory isolation:** To prevent transmission of infectious diseases over short distances through the air, a private room is indicated but patients infected with the same organism may share a room. In addition to the basic requirements, masks are indicated for those who come in close contact with the patient; gowns and gloves are not indicated.
 - d. **Tuberculosis isolation (AFB isolation):** For patients with pulmonary tuberculosis who have a positive sputum smear or chest-x-rays that strongly suggest active tuberculosis. Specifications include use of a private room with special ventilation and the door closed. In addition to the basic requirements, masks are used only if the patient is coughing and does not reliably and consistently cover the mouth. Gowns are used to prevent gross contamination of clothing. Gloves are not indicated.
 - e. **Enteric precautions:** For infections transmitted by direct or indirect contact with feces. In addition to the basic requirements, specifications include use of a private room if patient hygiene is poor. Masks are not indicated; gowns should be used if soiling is likely and gloves are to be used for touching contaminated materials.
 - f. **Drainage/secretion precautions:** To prevent infections transmitted by direct or indirect contact with purulent material or drainage from an infected body site. A private room and masking are not indicated; in addition to the basic requirements, gowns should be used if soiling is likely and gloves used for touching contaminated materials.

2503012423

g. **Blood/body fluid precautions:** To prevent infections that are transmitted by direct or indirect contact with infected blood or body fluids. In addition to the basic requirements, a private room is indicated if patient hygiene is poor; masks are not indicated; gowns should be used if soiling of clothing with blood or body fluids is likely. Gloves should be used for touching blood or body fluids.

See also QUARANTINE.

¹Benenson AS (Ed): *Control of Communicable Diseases in Man*, 14th ed. Washington DC: American Public Health Association, 1985.

ISOMETRIC CHART A chart or graph that portrays three dimensions on a plane surface.

J, K

JACKKNIFE A technique for estimating the variance and the bias of an estimator. If the sample size is n , the estimator is applied to each subsample of size $n-1$, obtained by dropping a measurement from analysis. The sum of squared differences between each of the resulting estimates and their mean, multiplied by $(n-1)/n$, is the jackknife estimate of variance; the difference between the mean and the original estimate, multiplied by $(n-1)$, is the jackknife estimate of bias.

JENNER, EDWARD (1749-1823) An English physician and naturalist. On the basis of the observation that dairymaids who had had cowpox never got smallpox, he inoculated a boy age 10 with cowpox (*vaccinia*) in 1796. Over the succeeding two years he inoculated 22 more persons and then attempted to inoculate them with smallpox, always without inducing this infection. The results of his work were published in *An Inquiry into the Cause and Effects of the Variola Vaccinae* (London, 1798). This successful method of immunizing persons and populations against smallpox led directly to the ultimate worldwide eradication of smallpox in 1977.

KAP (KNOWLEDGE, ATTITUDES, PRACTICE) SURVEY A formal survey, using face-to-face interviews, in which women are asked standardized pretested questions dealing with their knowledge of, attitudes toward, and use of contraceptive methods. Detailed reproductive histories and attitudes toward desired family size are also elicited. Analysis of responses provides much useful information on family planning and gives estimates of possible future trends in population structure. The term has sometimes been used to describe other varieties of survey of knowledge, attitudes, and practice, e.g., health promotion in general or in particular, cigarette smoking.

KAPPA A measure of the degree of nonrandom agreement between observers or measurements of the same categorical variable

$$k = \frac{P_o - P_e}{1 - P_e}$$

where P_o is the proportion of times the measurements agree, and P_e is the proportion of times they can be expected to agree by chance alone. If the measurements agree more often than expected by chance, kappa is positive; if concordance is complete, kappa = 1; if there is no more nor less than chance concordance, kappa = 0; if the measurements disagree more than expected by chance, kappa is negative.

KENDALL'S TAU See CORRELATION COEFFICIENT.

KOCH, ROBERT (1843-1910) German physician, pathologist, and bacteriologist. One of the founders of microbiology and an important contributor to our understanding of infectious disease epidemiology. His major contributions to medical science include the life cycle of anthrax, the etiology of traumatic infection, methods of fixing and staining bacteria, and, in 1882, the discovery of the tubercle bacillus. The paper

reporting this contained the first statement of KOCH'S POSTULATES. In 1883, he discovered the cholera vibrio. He was awarded the Nobel Prize in 1905.

KOCH'S POSTULATES First formulated by Henle and adapted by Robert Koch in 1877, with elaborations in 1882. Koch stated that these postulates should be met before a causative relationship can be accepted between a particular bacterial parasite or disease agent and the disease in question.

1. The agent must be shown to be present in every case of the disease by isolation in pure culture.
2. The agent must not be found in cases of other disease.
3. Once isolated, the agent must be capable of reproducing the disease in experimental animals.
4. The agent must be recovered from the experimental disease produced.

See also CAUSALITY: EVANS'S POSTULATES.

KURTOSIS The extent to which a unimodal distribution is peaked.

L

LARGE SAMPLE METHOD (Syn: asymptotic method): Any statistical method based on an approximation to a normal or other distribution that becomes more accurate as sample size increases. An example is a chi square test on a set of frequencies.

LATENT IMMUNIZATION The process of developing immunity by a single or repeated inapparent asymptomatic infection. Not necessarily related to latent infection. See also IMMUNITY, ACQUIRED.

LATENT INFECTION Persistence of an infectious agent within the host without symptoms (and often without demonstrable presence in blood, tissues, or bodily secretions of host).

LATENT PERIOD (Syn: latency) Delay between exposure to a disease-causing agent and the appearance of manifestations of the disease. After exposure to ionizing radiation, for instance, there is a latent period of five years, on average, before development of leukemia, and more than 20 years before development of certain other malignant conditions. The term "latent period" is often used synonymously with "induction period," that is, the period between exposure to a disease-causing agent and the appearance of manifestations of the disease. It has also been defined as the period from disease initiation to disease detection. See also INCUBATION PERIOD; INDUCTION PERIOD.

LATIN SQUARE One of the basic statistical designs for experiments that aim at removing from the experimental error the variation from two sources, which may be identified with the rows and columns of the square. In such a design the allocation of k experimental treatments in the cells of a k by k (latin) square is such that each treatment occurs exactly once in each row and column. A design for a 5×5 square is as follows:

| | | | | |
|---|---|---|---|---|
| A | B | C | D | E |
| B | A | E | C | D |
| C | D | A | E | B |
| D | E | B | A | C |
| E | C | D | B | A |

After Kendall and Buckland.¹

¹Kendall MG, Buckland AA: *A Dictionary of Statistical Terms*, 4th ed. London: Longman, 1982.
LAVERAN, ALPHONSE (1845–1922) French army surgeon who discovered the malaria parasite (1880) while on service in Algeria. Though initially sceptical, the scientific community soon accepted the validity of Laveran's discovery, which was confirmed and enlarged by Golgi, Grassi, and others. Laveran was awarded the Nobel Prize for medicine in 1907.

LEAD TIME The time gained in treating or controlling a disease when detection is earlier than usual, e.g., in the presymptomatic stage, as when screening procedures are used for detection.

LEAD TIME BIAS (Syn: zero time shift) Overestimation of survival time, due to the backward shift in the starting point for measuring survival that arises when diseases such as cancer are detected early, as by screening procedures.

LEAST SQUARES A principle of estimation, due to Gauss, in which the estimates of a set of parameters in a statistical model are those quantities that minimize the sum of squared differences between the observed values of the dependent variable and the values predicted by the model.

LEDERMANN FORMULA Ledermann¹ showed empirically that the frequency distribution of alcohol consumption in the population of consumers may be log-normal; the curve is sharply skewed—approximately one-third of drinkers consume more than 60% of the total amount of alcohol. Among drinkers the proportion of persons with alcoholism remains constant at around 7–9%. The pattern of consumption of illicit drugs among users may also be log-normal. Questions have been raised, however, about the validity of some assumptions upon which the formula is based.

¹Ledermann S: *Alcool, Alcoolisme et Alcoolisation*. Paris: Presses universitaires de France, 1956.

LEEUWENHOEK, ANTONI VAN (1632–1723) An early microscopist from Delft, in the Netherlands, the first to use his microscopes to examine and describe small creatures (*animalcules*) such as the protozoan organisms in vaginal secretions, spermatozoa, and with growing ability to make more powerful microscopes, infectious microorganism. He was thus a key figure in the development of the germ theory of disease.

LEVIN'S ATTRIBUTABLE RISK See ATTRIBUTABLE FRACTION (POPULATION).

LIFE EVENTS Changes or disruptions in the pattern of living that may be associated with or produce changes in health. The relationship of "life stress" and "emotional stress" to onset of several kinds of serious chronic disease such as coronary heart disease and hypertension has been the subject of epidemiologic studies. The Rahe–Holmes Social Readjustment Rating Scale¹ was the first to be developed to assign ranks or ratings to significant life events such as death of a spouse or other close relative, loss of regular job, relocation, marriage, divorce, etc. Many other rating scales have since been developed.

¹Holmes TH, Rahe RH: The social readjustment rating scale. *J Psychosomatic Res* 1:213–218, 1967.

LIFE EXPECTANCY See EXPECTATION OF LIFE.

LIFE EXPECTANCY FREE FROM DISABILITY (LEFD) An estimate of life expectancy adjusted for activity-limitation (data for which are derived from hospital discharge statistics, etc.). See also QALY.

LIFE STYLE The set of habits and customs that is influenced, modified, encouraged, or constrained by the lifelong process of socialization. These habits and customs include use of substances such as alcohol, tobacco, tea, coffee; dietary habits, exercise, etc., which have important implications for health and are often the subject of epidemiologic investigations.

LIFE TABLE A summarizing technique used to describe the pattern of mortality and survival in populations. The survival data are time specific and cumulative probabilities of survival of a group of individuals subject, throughout life, to the age-specific death rates in question. The life table method can be applied to the study not only of death, but also of any defined endpoint such as the onset of disease or the occurrence of specific complication(s) of disease. The survivors to age x are denoted by the symbol l_x , the expectation of life at age x is denoted by the symbol

t_n , and the proportion alive at age x who die between age x and $x+1$ years is denoted by the symbol nq_x . The life table method is used extensively in epidemiology and in many assessments of treatment regimens in clinical practice.

The first rudimentary life tables were published in 1693 by the astronomer Edmund Halley. These made use of records of the funerals in the city of Breslau. In 1815 in England, the first actuarially correct life table was published, based on both population and death data classified by age.

Two types of life tables may be distinguished according to the reference year of the table: the current or period life table and the generation or cohort life table.

The current life table is a summary of mortality experience over a brief period (one to three years), and the population data relate to the middle of that period (usually close to the date of a census). A current life table therefore represents the combined mortality experience by age of the population in a particular short period of time.

The cohort or generation life table describes the actual survival experience of a group, or cohort, of individuals born at about the same time. Theoretically, the mortality experience of the persons in the cohort would be observed from their moment of birth through each consecutive age in successive calendar years until all of them die.

The clinical life table describes the outcome experience of a group or cohort of individuals classified according to their exposure or treatment history.

Life tables are also classified according to the length of age interval in which the data are presented. A complete life table contains data for every single year of age from birth to the last applicable age. An abridged life table contains data by intervals of five or ten years of age. See also EXPECTATION OF LIFE: SURVIVORSHIP STUDY.

LIFE TABLE, EXPECTATION OF LIFE FUNCTION, \bar{t}_x (Syn: average future lifetime) The expectation of life function is a statement of the average number of years of life remaining to persons who survive to age x .

LIFE TABLE, SURVIVORSHIP FUNCTION, I_x The survivorship function is a statement of the number of persons out of an initial population of defined size, e.g., 100,000 live births, who would survive or remain free of a defined endpoint condition to age x under the age-specific rates for the specified year. The value of I_{x_0} , for example, is determined by the cumulative operation of the specific death rates for all ages below 40.

LIFETIME RISK The risk to an individual that a given health effect will occur at any time after exposure, without regard for the time at which that effect occurs.

LIKELIHOOD FUNCTION A function constructed from a statistical model and a set of observed data, which gives the probability of the observed data for various values of the unknown model parameters. The parameter values that maximize the probability are the maximum likelihood estimates of the parameters.

LIKELIHOOD RATIO TEST A statistical test based on the ratio of the maximum value of the likelihood function under one statistical model to the maximum value under another statistical model; the models differ in that one includes, the other excludes, one or more parameters.

LIND, JAMES (1716-1794) British naval surgeon; contributed to improved hygiene aboard ships. Conducted what amounted to epidemiologic experiments (albeit with small numbers) which established that scurvy could be prevented by fresh fruits such as lemons and oranges.

LINEAR MODEL A statistical model in which the value of a parameter for a given value of a factor, x , is assumed to be equal to $a + bx$, where a and b are constants.

LINEAR REGRESSION Regression analysis of data using linear models.

LINKAGE See GENETIC LINKAGE; RECORD LINKAGE.

LIVE BIRTH WHO definition adopted by Third World Health Assembly, 1950: Live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered live born.

In the *Report of WHO Expert Committee on Prevention of Perinatal Mortality and Morbidity (Technical Report Series 457, 1970)*, it is noted that the above definition requires the inclusion as live births of very early and patently nonviable fetuses and that accordingly it is not strictly applied. The committee suggested, therefore, that WHO should introduce a viability criterion into the definition so that very immature fetuses surviving for very short periods were excluded, even though they showed one or more of the transitory signs of life.

LOCUS

1. The position of a point, as defined by the coordinates on a graph.
2. The position that a gene occupies on a chromosome.

LOG SCORE In genetics, the log odds ratio of observed to expected distribution of genetic markers.

LOGISTIC MODEL A statistical model of an individual's risk (probability of disease y) as a function of a risk factor x :

$$P(y|x) = \frac{1}{1 + e^{-a - bx}}$$

where e is the (natural) exponential function. This model has a desirable range, 0 to 1, and other attractive statistical features. In the multiple logistic model, the term bx is replaced by a linear term involving several factors, e.g., $\beta_1x_1 + \beta_2x_2$ if there are two factors x_1 and x_2 .

LOGIT (Syn: log-odds) The logarithm of the ratio of frequencies of two different categorical outcomes such as healthy versus sick.

LOGIT MODEL A linear model for the logit (natural log of the odds) of disease as a function of a quantitative factor X :

$$\text{Logit (disease given } X=x) = a + bx$$

This model is mathematically equivalent to the LOGISTIC MODEL.

LOG-LINEAR MODEL A statistical model that uses an ANALYSIS OF VARIANCE type of approach for the modeling of frequency counts in contingency tables.

LOG-NORMAL DISTRIBUTION If a variable Y is such that $X = \log Y$ is normally distributed, it is said to have log-normal distribution. This is a SKEW DISTRIBUTION. See also NORMAL DISTRIBUTION.

LONGITUDINAL STUDY See COHORT STUDY.

Louis, PIERRE-CHARLES-ALEXANDRE (1787-1872) French physician and mathematician. One of the founders of medical statistics, his research on tuberculosis, which included dissection of 358 specimens and study of 1960 clinical cases, led to publication of *Recherches anatomicopathologiques sur la phthisie* (Paris, 1825). This work and others are marked by rigorous numerical precision and demonstration of similarities and differences based upon numerical distribution of data. The Lilienfelds'

have pointed out that Louis greatly influenced the development of statistics as applied in biology and medicine; he either taught or otherwise directly influenced many European, British, and American workers, including William Farr, John Simon, William Augustus Guy, and William Budd in England, George Shattuck, Elisha Barnett, and Alonzo Clark in the United States, and Joseph Skoda in Hungary; those he influenced handed on these important concepts to their own pupils.

¹Lilienfeld AM, Lilienfeld D: Threads of epidemiological history, in *Foundations of Epidemiology*, 2nd Ed. (New York: Oxford, 1980), pp. 23-45.

LOW BIRTH WEIGHT See **BIRTH WEIGHT**.

"LUMPING AND SPLITTING" Derisive term describing the propensity of epidemiologists to group related phenomena or to separate phenomena that hitherto have been grouped. Epidemiologists are sometimes called "lumpers and splitters."

M

MALARIA ENDEMICITY Certain terms used to describe the occurrence of malaria, based on enlarged spleen rates are categorized by WHO as follows:

1. Hypoendemic: Spleen rate in children 2-9 years <10%.
2. Mesoendemic: Spleen rate 11-50%.
3. Hyperendemic: Spleen rate in children over 50%, in adults usually over 25%.
4. Holoendemic: Spleen rate in children constantly over 75%, adult rate low.

MALARIA PERIODICITY Recurrence at regular intervals of symptoms; periodicity may be quotidian, tertian, or quartan, according to the interval between paroxysms:

1. Quartan: Recurring every third day, i.e., day 1, day 4, day 7, etc.
2. Quotidian: Recurring daily.
3. Tertian: Recurring every alternate day, i.e., day 1, day 3 etc.

MALARIA PATENT PERIOD Period during which parasites are present in peripheral blood.

MALARIA REPRODUCTION RATE Estimated number of malarial infections potentially distributed by the average nonimmune infected individual in a community where neither persons nor mosquitoes were previously infected.

MALARIA SURVEY Investigation in selected age-group samples in randomly selected localities to assess malaria endemicity; uses spleen and/or parasite rates as measure of endemicity.

MALTHUS, THOMAS ROBERT (1766-1834) An English clergyman and natural scientist who argued in *An Essay on the Principle of Population* (London, 1798) that populations increase in geometric progression while food supplies increase only in arithmetical progression, thus making famine inevitable. His work justifies his recognition as one of the founders of demography, even though events proved his predictions wrong (at least in the short term).

MANSON, PATRICK (1844-1922) Studied tropical diseases in China and made many contributions of fundamental importance, notably the transmission of filariasis by culicine mosquitoes, parts of the life cycle of schistosomes. He investigated and observed many other tropical parasitic diseases and founded the London School of (Hygiene and) Tropical Medicine in 1898.

MANTEL-HAENSZEL ESTIMATE, MANTEL-HAENSZEL ODDS RATIO Mantel and Haenszel¹ provided an adjusted odds ratio as an estimate of relative risk that may be derived from grouped and matched sets of data. It is now known as the Mantel-Haenszel estimate, one of the few eponymous terms of modern epidemiology.

The statistic may be regarded as a type of weighted average of the individual odds ratios, derived from stratifying a sample into a series of strata that are internally homogeneous with respect to confounding factors.

The Mantel-Haenszel summarization method can also be extended to the summarization of rate ratios and rate differences from follow-up studies.

¹Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22:719-748, 1959.

2503012427

MANTEL-HAENSZEL TEST A summary CHI-SQUARE TEST developed by Mantel and Haenszel for stratified data and used when controlling for CONFOUNDING.

MARGIN OF SAFETY An estimate of the ratio of the no-observed-effect level (NOEL) to the level accepted in regulations.

MARGINALS The row and column totals of a contingency table.

MARKOV PROCESS A stochastic process such that the conditional probability distribution for the state at any future instant, given the present state, is unaffected by any additional knowledge of the past history of the system.

MASKED STUDY See BLINDED STUDY.

MASKING Procedure(s) intended to keep participant(s) in a study from knowing some fact(s) or observation(s) that might bias or influence their actions or decisions regarding the study.

MATCHED CONTROLS See CONTROLS, MATCHED.

MATCHING The process of making a study group and a comparison group comparable with respect to extraneous factors. Several kinds of matching can be distinguished:

Caliper matching is the process of matching comparison group subjects to study group subjects within a specified distance for a continuous variable (e.g., matching age to within two years).

Frequency matching requires that the frequency distributions of the matched variable(s) be similar in study and comparison groups.

Category matching is the process of matching study and control group subjects in broad classes such as relatively wide age ranges or occupational groups.

Individual matching relies on identifying individual subjects for comparison, each resembling a study subject on the matched variable(s).

Pair matching is individual matching in which study and comparison subjects are paired.

MATERNAL MORTALITY (RATE) The risk of dying from causes associated with childbirth is measured by the maternal mortality rate. For this purpose the deaths used in the numerator are those arising during pregnancy or from puerperal causes, i.e., deaths occurring during and/or due to deliveries, complications of pregnancy, childbirth, and the puerperium. Women exposed to the risk of dying from puerperal causes are those who have been pregnant during the period. Their number being unknown, the number of live births is used as the conventional denominator for computing comparable maternal mortality rates. The formula is

$$\text{Annual maternal mortality rate} = \frac{\text{Number of deaths from puerperal causes in a given geographic area during a given year}}{\text{Number of live births that occurred among the population of the given geographic area during the same year}} \times 1000 \text{ (or } 100,000\text{)}$$

There is variation in the duration of the postpartum period in which death may occur and be certified due to "puerperal causes," i.e., "maternal mortality." According to WHO, a maternal death is defined as the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.

Maternal deaths should be subdivided into two groups: (1) direct obstetric deaths,

12503012428

resulting from obstetric complications of the pregnant state, and (2) indirect obstetric deaths, resulting from preexisting disease or conditions not due to direct obstetric causes.

Although WHO defines maternal mortality as death during pregnancy or within 42 days of delivery, in some jurisdictions, a period as long as a year is used.

MATHEMATICAL MODEL A representation of a system, process, or relationship in mathematical form in which equations are used to simulate the behavior of the system or process under study. The model usually consists of two parts: the mathematical structure itself, e.g., Newton's inverse square law or Gauss's "normal" law, and the particular constants or parameters associated with them, such as Newton's gravitational constant or the Gaussian standard deviation.

A mathematical model is deterministic if the relations between the variables involved take on values not allowing for any play of chance. A model is said to be statistical, stochastic, or random, if random variation is allowed to enter the picture. See also MODEL.

MAXIMUM ALLOWABLE CONCENTRATION (MAC) See SAFETY STANDARDS.

MAXIMUM LIKELIHOOD ESTIMATE The value for an unknown parameter that maximizes the probability of obtaining exactly the data that were observed.

MCNEMAR'S TEST A form of the CHI-SQUARE TEST for matched-pairs data. It is a special case of the MANTEL-HAENSZEL TEST.

MEAN, ARITHMETIC A MEASURE OF CENTRAL TENDENCY. Calculable only for positive values. It is calculated by taking the logarithms of the values, calculating their arithmetic mean, then converting back by taking the antilogarithm.

MEAN, HARMONIC A MEASURE OF CENTRAL TENDENCY computed by summing the reciprocals of all the individual values and dividing the resulting sum into the number of values.

MEASURE OF ASSOCIATION A quantity that expresses the strength of association between variables. Commonly used measures of association are differences between means, proportions or rates, the rate ratio, the odds ratio, and correlation and regression coefficients.

MEASUREMENT The procedure of applying a standard scale to a variable or to a set of values.

MEASUREMENT, PROBLEMS WITH TERMINOLOGY There is sometimes uncertainty about the terms used to describe the properties of measurement: accuracy, precision, validity, reliability, repeatability, and reproducibility. Accuracy and precision are often used synonymously, validity is defined variously, and reliability, repeatability, and reproducibility are frequently used interchangeable.

Etymologies are helpful in making a case for preferred usages, but they are not always decisive. Accuracy is from the Latin *cura*, care, and while this may be of interest to those in the health field, it does not illuminate the origins of the standard definition, that is, "conforming to a standard or a true value" (OED). Accuracy is distinguished from precision in this way: A measurement or statement can reflect or represent a true value without detail. A temperature reading of 98.6°F is accurate, but it is not precise if a more refined thermometer registers a temperature of 98.637°F.

Precision (from Latin *præcidiere*, cut short) is the quality of being sharply defined through exact detail. A faulty measurement may be expressed precisely, but may not be accurate. Measurements should be both accurate and precise, but the two terms are not synonymous. Consistency or reliability describes the property of measurements or results that conform to themselves.

Reliability (Latin *reliare*, to bind) is defined by the *OED* as a quality that is sound and dependable. Its epidemiologic usage is similar; a result or measurement is said to be reliable when it is stable, i.e., when repetition of an experiment or measurement gives the same results. The terms "repeatability" and "reproducibility" are synonymous (the *OED* defines each in terms of the other), but they do not refer to a quality of measurement, rather only to the action of performing something more than once. Thus, a way of discovering whether or not a measurement is reliable is to repeat or reproduce it. The terms "repeatability" and "reproducibility," formed from their respective verbs, are used inaccurately when they are substituted for "reliability," a noun that refers to the measuring procedure rather than the attribute being measured. However, in common usage, both repeatability and reproducibility refer to the capacity of a measuring procedure to produce the same result on each occasion in a series of procedures conducted under identical conditions.

Validity is used correctly when it agrees with the standard definition given by the *OED*: "sound and sufficient." If, in the epidemiologic sense, a test measures what it purports to measure (it is sufficient) then the test is said to be valid. See also ACCURACY; PRECISION; RELIABILITY; REPEATABILITY; VALIDITY.

MEASUREMENT SCALE The complete range of possible values for a measurement (e.g., the set of possible responses to a question, the physically possible range for a set of body weights). Measurement scales are sometimes classified into five major types, according to the quantitative character of the scale:

1. **Dichotomous scale:** One that arranges items into either of two mutually exclusive categories.
2. **Nominal scale:** Classification into unordered qualitative categories; e.g., race, religion, and country of birth as measurements of individual attributes are purely nominal scales, as there is no inherent order to their categories.
3. **Ordinal scale:** Classification into ordered qualitative categories, e.g., social class (I, II, III etc.), where the values have a distinct order, but their categories are qualitative in that there is no natural (numerical) distance between their possible values.
4. **Interval scale:** An (equal) interval involves assignment of values with a natural distance between them, so that a particular distance (interval) between two values in one region of the scale meaningfully represents the same distance between two values in another region of the scale. Examples include Celsius and Fahrenheit temperature, date of birth.
5. **Ratio scale:** A ratio is an interval scale with a true zero point, so that ratios between values are meaningfully defined. Examples are absolute temperature, weight, height, blood count, and income, as in each case it is meaningful to speak of one value as being so many times greater or less than another value.

MEASURES OF CENTRAL TENDENCY A general term for several characteristics of the distribution of a set of values or measurements around a value or values at or near the middle of the set. The principal measures of central tendency are the mean (average), median, and mode. (See entries under each.)

MECHANICAL TRANSMISSION See VECTOR-BORNE INFECTION.

MEDIAN A MEASURE OF CENTRAL TENDENCY. The simplest division of a set of measurements is into two parts—the lower and the upper half. The point on the scale that divides the group in this way is called the "median."

MEDIATOR (MEDIATING) VARIABLE See INTERMEDIATE VARIABLE.

MEDICAL AUDIT A health service evaluation procedure in which selected data from pa-

tients' charts are summarized in tables displaying such data as average length of stay or duration of an episode of care, the frequency of diagnostic and therapeutic procedures, and outcomes of care arranged by diagnostic category. These are often compared with predetermined norms.

MEDICAL CARE See HEALTH CARE.

MEDICAL RECORD A file of information relating to transaction(s) in personal health care. In addition to facts about a patient's illness, medical records nearly always contain other information. The full range of data in medical records includes the following:

1. Clinical, i.e., diagnosis, treatment, progress, etc.
2. Demographic, i.e., age, sex, birthplace, residence, etc.
3. Sociocultural, i.e., language, ethnic origin, religion, etc.
4. Sociological, i.e., family (next of kin), occupation, etc.
5. Economic, i.e., method of payment (fee-for service, indigent, etc.).
6. Administrative, i.e., site of care, provider, etc.
7. "Behavioral," e.g., record of broken appointment may indicate dissatisfaction with service provided.

MEDICAL STATISTICS See BIOSTATISTICS.

MENDEL'S LAWS Derived from the pioneering genetic studies of Gregor Mendel (1822–1884). Mendel's first law states that genes are particulate units that segregate; i.e., members of the same pair of genes are never present in the same gamete, but always separate and pass to different gametes. Mendel's second law states that genes assort independently; i.e., members of different pairs of genes move to gametes independently of one another.

META-ANALYSIS The process of using statistical methods to combine the results of different studies. In the biomedical sciences, the systematic, organized and structured evaluation of a problem of interest, using information (commonly in the form of statistical tables or other data) from a number of independent studies of the problem. A frequent application has been the pooling of results from a number of small randomized controlled trials, none in itself large enough to demonstrate statistically significant differences, but in aggregate, capable of so doing. Meta-analysis has a qualitative component, i.e., application of predetermined criteria of quality (e.g., completeness of data, absence of biases), and a quantitative component, i.e., integration of the numerical information. Meta-analysis includes aspects of an overview, and of pooling of data, but implies more than either of these processes. Meta-analysis carries the risk of several biases.

METHODOLOGY The scientific study of methods. Methodology should not be confused with methods. Sad to say, the word "methodology" is all too often used when the writer means "method."

MIASMA THEORY An explanation for the origin of epidemics, the "miasma theory" was implied by many ancient writers, and made explicit by Lancisi in *De noxiis paludum effervesciis* (1717). It was based on the notion that when the air was of a "bad quality" (a state that was not precisely defined, but that was supposedly due to decaying organic matter), the persons breathing that air would become ill. Malaria ("bad air") is the classic example of a disease that was long attributed to miasms. "Miasma" was believed to pass from cases to susceptibles in these diseases considered contagious.

MIGRANT STUDIES Studies taking advantage of migration to one country by those from other countries with different physical and biological environments, cultural background and/or genetic makeup, and different morbidity or mortality experience. Comparisons are made between the mortality or morbidity experience of the mi-

grant groups with that of their current country of residence and/or their country of origin. Sometimes the experiences of a number of different groups who have migrated to the same country have been compared.

MILL'S CANONS In *A System of Logic* (1856), J.S. Mill devised logical strategies (canons) from which causal relationships may be inferred. Four in particular are pertinent to epidemiology: the methods of agreement, difference, residues, and concomitant variation.

Method of agreement (first canon): "If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon."

Method of difference (second canon): "If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former, the circumstance in which alone the two instances differ is the effect, or cause or a necessary part of the cause, of the phenomenon."

Method of residues (fourth canon): "Subduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect of the remaining antecedents."

Method of concomitant variation (fifth canon): "Whatever phenomenon varies in any manner whether another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation."

MINIMUM DATA SET (Syn: uniform basic data set) A widely agreed upon and generally accepted set of terms and definitions constituting a core of data acquired for medical records and employed for developing statistics suitable for diverse types of analyses and users. Such sets have been developed for birth and death certificates, ambulatory care, hospital care, and long-term care. See also **BIRTH CERTIFICATE**; **DEATH CERTIFICATE**; **HOSPITAL DISCHARGE ABSTRACT SYSTEM**.

MISCLASSIFICATION The erroneous classification of an individual, a value, or an attribute into a category other than that to which it should be assigned. The probability of misclassification may be the same in all study groups (nondifferential misclassification) or may vary between groups (differential misclassification).

MOBILITY, GEOGRAPHIC Movement of persons from one country or region to another.

MOBILITY, SOCIAL Movement from one defined socioeconomic group to another, either upward or downward. Downward social mobility, which can be related to impaired health (e.g., alcoholism, schizophrenia, or mental retardation), is sometimes referred to as "social drift."

MODE One of the **MEASURES OF CENTRAL TENDENCY**. The most frequently occurring value in a set of observations.

MODEL

1. An abstract representation of the relationship between logical, analytical, or empirical components of a system. See also **MATHEMATICAL MODEL**.
2. A formalized expression of a theory or the causal situation that is regarded as having generated observed data.
3. (Animal) model: an experimental system that uses animals, because humans cannot be used for ethical or other reasons.
4. A small-scale simulation, e.g., by using an "average region" with characteristics resembling those of the whole country.

2503012430

In epidemiology the use of models began with an effort to predict the onset and course of epidemics. In the second report of the Registrar-General of England and Wales (1840), WILLIAM FARR developed the beginnings of a predictive model for communicable disease epidemics. He had recognized regularities in the smallpox epidemics of the 1830s. By calculating frequency curves for these past outbreaks, he estimated the deaths to be expected. See also **DEMONSTRATION MODEL**; **MATHEMATICAL MODEL**; **THEORETICAL EPIDEMIOLOGY**.

MODERATOR VARIABLE (Syn: qualifier variable) In a study of a possible causal factor and an outcome, a moderator variable is a third variable exhibiting statistical interaction by virtue of its being antecedent or intermediate in the causal process under study. If it is antecedent, it is termed a conditional moderator variable or **EFFECT MODIFIER**; if it is intermediate, it is a contingent moderator variable. See also **INTERACTION**; **INTERMEDIATE VARIABLE**.

MONITORING

1. The performance and analysis of routine measurements, aimed at detecting changes in the environment or health status of populations. Not to be confused with **SURVEILLANCE**. To some, monitoring also implies intervention in the light of observed measurements.
2. Ongoing measurement of performance of a health service or a health professional, or of the extent to which patients comply with or adhere to advice from health professionals.
3. In management, the continuous oversight of the implementation of an activity that seeks to ensure that input deliveries, work schedules, targeted outputs, and other required actions are proceeding according to plan.

MONOTONIC SEQUENCE A sequence is said to be monotonic increasing if each value is greater than or equal to the previous one, and monotonic decreasing if each value is less than or equal to the previous one. If equality of values is excluded, we speak of a strictly (increasing or decreasing) monotonic sequence.

MONTE CARLO STUDY, TRIAL Complex relationships that are difficult to solve by mathematical analysis are sometimes studied by computer experiments that simulate and analyze a sequence of events, using random numbers. Such experiments are called Monte Carlo trials, or studies, in recognition of Monte Carlo as one of the gambling capitals of the world.

MORBIDITY Any departure, subjective or objective, from a state of physiological or psychological well-being. In this sense, **sickness**, **illness**, and **morbid condition** are similarly defined and synonymous (but see **DISEASE**).

The WHO Expert Committee on Health Statistics noted in its Sixth report (1959) that morbidity could be measured in terms of three units: (1) persons who were ill; (2) the illnesses (periods or spells of illness) that these persons experienced; and (3) the duration (days, weeks, etc.) of these illnesses. See also **HEALTH INDEX**; **INCIDENCE RATE**; **NOTIFIABLE DISEASE**; **PREVALENCE RATE**.

MORBIDITY RATE A term, preferably avoided, used indiscriminately to refer to incidence or prevalence rates of disease.

MORBIDITY SURVEY A method for estimating the prevalence and/or incidence of disease or diseases in a population. A morbidity survey is usually designed simply to ascertain the facts as to disease distribution, and not to test a hypothesis. See also **CROSS-SECTIONAL STUDY**; **HEALTH SURVEY**.

MORTALITY RATE See **DEATH RATE**.

MORTALITY STATISTICS Statistical tables compiled from the information contained in **DEATH**

CERTIFICATES. Most administrative jurisdictions in all nations produce tables of mortality statistics. These may be published at regular intervals; they usually show numbers of deaths and/or rates by age, sex, cause, and sometimes other variables.

MULTICOLLINEARITY In multiple regression analysis, a situation in which at least some of the independent variables are highly correlated with each other. Such a situation can result in inaccurate estimates of the parameters in the regression model.

MULTIFACTORIAL ETIOLOGY See MULTIPLE CAUSATION.

MULTINOMIAL DISTRIBUTION The probability distribution associated with the classification of each of a sample of individuals into one of several mutually exclusive and exhaustive categories. When the number of categories is two, the distribution is called binomial. See also BINOMIAL DISTRIBUTION.

MULTIPHASIC SCREENING See SCREENING.

MULTIPLE CAUSATION (Syn: multifactorial etiology) This term is used to refer to the concept that a given disease or other outcome may have more than one cause. A combination of causes or alternative combinations of causes may be required to produce the effect.

MULTIPLE LOGISTIC MODEL See LOGISTIC MODEL.

MULTIPLE RISK Where more than one risk factor for the development of a disease or other outcome is present, and their combined presence results in an increased risk, we speak of "multiple risk." The increased risk may be due to the additive effects of the risks associated with the separate risk factors, or to SYNERGISM.

MULTIPLICATIVE MODEL A model in which the joint effect of two or more causes is the product of their effects. For instance, if factor a multiplies risk by the amount a in the absence of factor b , and factor b multiplies risk by the amount b in the absence of factor a , the combined effect of factors a and b on risk is $a \times b$. See also ADDITIVE MODEL.

MULTISTAGE MODEL A mathematical model, mainly for carcinogenesis, based on the theory that a specific carcinogen may affect one of a number of stages in the development of cancer.

MULTIVARIATE ANALYSIS A set of techniques used when the variation in several variables has to be studied simultaneously. In statistics, any analytic method that allows the simultaneous study of two or more DEPENDENT VARIABLES.

MUTATION Heritable change in the genetic material not caused by genetic segregation or recombination, which is transmitted to daughter cells and to succeeding generations, provided it is not a dominant lethal factor.

MUTATION RATE The frequency with which mutations occur per gene or per generation.

N

NATIONAL DEATH INDEX A computerized central registry of deaths in the United States, started in 1979 and operated by the U.S. National Center for Health Statistics, that facilitates mortality followup; cf. CANADIAN MORTALITY DATA BASE.

NATURAL EXPERIMENT A term probably derived from JOHN SNOW's account of his investigation of the practices of water supply companies in relation to the cholera epidemics in London in the 1850s. It refers to naturally occurring circumstances in which populations have different exposures to a supposed causal factor in a situation resembling an actual experiment in which persons would be assigned to groups.

John Snow was able to trace the London outbreaks of cholera in the 19th century to water impurity as a result of comparisons made between two water companies. It would have been unethical to expose "test subjects" to infection, but the situation at the time afforded him the opportunity to make observations of crucial importance.

To turn this grand experiment to account, all that was required was to learn the supply of water to each individual house where a fatal attack of cholera might occur . . . I resolved to spare no exertion which might be necessary to ascertain the exact effect of the water supply on the progress of the epidemic, in the places where all the circumstances were so happily adapted for the inquiry . . . I had no reason to doubt the correctness of the conclusions I had drawn from the great number of facts already in my possession, but I felt that the circumstances of the cholera-poisoning passing down the sewers into a great river, and being distributed through miles of pipes, and yet producing its specific effects was a fact of so startling a nature, and of so vast importance to the community, that it could not be too rigidly examined or established on too firm a basis. (Snow, *On the Mode of Communication of Cholera*, 1855)

NATURAL HISTORY OF DISEASE The course of a disease from onset (inception) to resolution. Many diseases have certain well-defined stages that, taken all together, are referred to as the "natural history of the disease" in question. These stages are as follows:

1. Stage of pathological onset.
2. Presymptomatic stage: from onset to the first appearance of symptoms and/or signs. SCREENING tests may lead to earlier detection.
3. Clinically manifest disease, which may progress inexorably to a fatal termination, be subject to remissions and relapses, or regress spontaneously, leading to recovery.

Detection and intervention can alter the natural history of disease. The term has also been used to mean "descriptive epidemiology of disease."

NATURAL HISTORY STUDY A study, generally longitudinal, designed to yield information about the natural course of a disease or condition.

NATURAL RATE OF INCREASE (DECREASE) See GROWTH RATE OF POPULATION.

NEAREST NEIGHBOR METHOD A means of analyzing the spatial patterns of a free-living population. A term from veterinary epidemiology. Random sampling points are located throughout an area and the distance from each point to the nearest individual is measured; alternatively, individuals are selected at random and from each of these the distance to the nearest neighbor is measured.

18741203012431

NECESSARY AND SUFFICIENT CAUSE A causal factor whose presence is required for the occurrence of the effect and whose presence is always followed by the effect. See also **ASSOCIATION; CAUSALITY**.

NEEDS (Syn: health needs, perceived needs, professionally defined needs, unmet needs) This term has both a precise and an all-but-indefinable meaning in the context of public health. We speak of needs in precise numerical terms when we refer to specific indicators of disease or premature death that require intervention because their level is above that generally accepted in the society or community in question. For example, an infant mortality rate two or three times greater than the national average in a particular community is an indicator of unmet health needs of infants in that community (not to be confused with a need for more or better medical care). It should be clear that even in this seemingly precise usage there are implied value judgments. It must be explicitly stated that "needs" always reflect prevailing value judgments as well as the existing ability to control a particular public health problem. Thus, sputum-positive pulmonary tuberculosis was not recognized as a health need in 1850 but was by 1900 in the industrialized nations; the ill effects of cigarette smoking must now be universally acknowledged as a health need; and child abuse is increasingly regarded as a public health problem, to which we could apply the term "professionally defined need."

(See Vickers GR: What sets the goals of public health. *Lancet* 1:599, 1958.)

NEONATAL MORTALITY RATE

1. In **VITAL STATISTICS**, the number of deaths in infants under 28 days of age in a given period, usually a year, per 1000 live births in that period.
2. In obstetric and perinatal research the term "neonatal mortality rate" is often used to denote the cumulative **MORTALITY RATE** of live-born infants within 28 days of age.

NESTED CASE CONTROL STUDY A case control study in which cases and controls are drawn from the population in a cohort study. As some data are already available about both cases and controls, the effects of some potential confounding variables are reduced or eliminated.

NET MIGRATION The numerical difference between immigration and emigration.

NET MIGRATION RATE The net effect of immigration and emigration on an area's population expressed as an increase or decrease per 1000 population of the area in a given year.

NET REPRODUCTION RATE The average number of female children born per woman in a cohort subject to a given set of age-specific fertility rates, a given set of age-specific mortality rates, and a given sex ratio at birth. This rate measures replacement fertility under given conditions of fertility and mortality: it is the ratio of daughters to mothers assuming continuation of the specified conditions of fertility and mortality. It is a measure of population growth from one generation to another under constant conditions. This rate is similar to the gross reproduction rate, but takes into account that some women will die before completing their childbearing years. An NRR of 1.00 means that each generation of mothers is having exactly enough daughters to replace itself in the population. See also **GROSS REPRODUCTION RATE**.

NEW YORK STATE IDENTIFICATION AND INTELLIGENCE SYSTEM (NYSIIS) A method of identifying individuals for **RECORD LINKAGE** based on phonetic spelling of full names, sequence of digits for birthdate, birthplace, sex, name at birth, and parents' names. See also **HOGBEN NUMBER; SOUNDEX CODE**.

NIDUS A focus of infection. The term can be used to describe any heterogeneity in the distribution of a disease, but is usually applied to a small area in which conditions

favor occurrence and spread of a communicable disease; also, the site of origin of a pathological process.

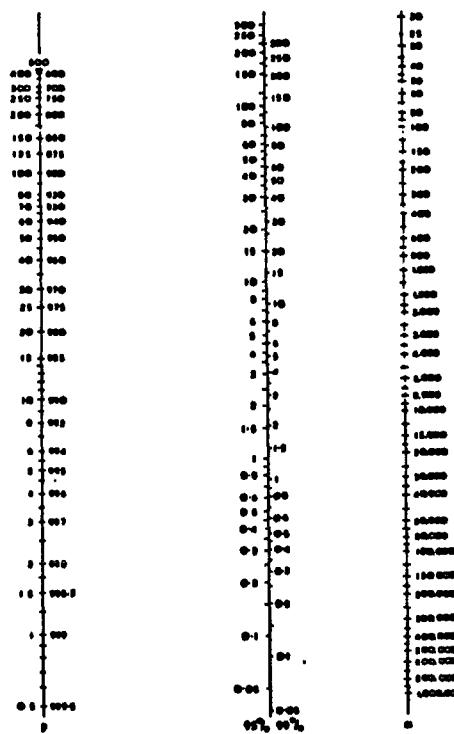
NIGHTINGALE, FLORENCE (1820-1910) An English woman who is identified as the founder of modern nursing, but was much more. In addition to her famous work of elevating nursing to a noble profession during the Crimean War, and establishing a training school for nurses at St. Thomas's Hospital in London, she recognized the importance of statistical analysis of hospital records (*Notes on Hospitals* London: Longmans, 1859); her contributions were recognized by election to Fellowship of the Royal Statistical Society. Her best-known work is *Notes on Nursing* (1860).

NOISE (IN DATA) This term is used when extraneous uncontrolled variables and/or errors influence the distribution of measurements that are made in a study, thus rendering difficult or impossible the determination of relationships between variables under scrutiny.

NOMENCLATURE A list of all approved terms for describing and recording observations.

NOMINAL SCALE See **MEASUREMENT SCALE**.

NOMOGRAM A form of line chart showing scales for the variables involved in a particular formula in such a way that corresponding values for each variable lie on a straight line intersecting all the scales.



Nomogram of confidence limits to a rate.
From Rosenbaum, Nomograms for rates per 1000, *Br Med J* 1:169-170, 1963.

NONCONCURRENT STUDY See HISTORICAL COHORT STUDY.
NONDIFFERENTIAL MISCLASSIFICATION See MISCLASSIFICATION.

NONEXPERIMENTAL STUDY See OBSERVATIONAL STUDY.

NONPARAMETRIC METHODS See DISTRIBUTION-FREE METHOD.

NONPARAMETRIC TEST See DISTRIBUTION-FREE METHOD.

NONPARTICIPANTS (Syn: nonresponders) Members of a study sample or population who do not take part in the study for whatever reason, or members of a target population who do not participate in an activity. Differences between participants and nonparticipants have been demonstrated repeatedly in studies of many kinds, and this is often a source of BIAS.

NO-OBSERVED-EFFECT LEVEL (NOEL) A term from toxicology, meaning the highest dose at which no adverse health effects are detected in an animal population. A NOEL-SF is a no-observed-effects level with an added safety factor for human exposures, used in setting human safety standards.

NORM This term has two quite distinct meanings:

1. The first is "what is usual," e.g., the range into which blood pressure values usually fall in a population group, the dietary or infant feeding practices that are usual in a given culture, or the way that a given illness is usually treated in a given health care system.
2. The second sense is "what is desirable," e.g., the range of blood pressures that a given authority regards as being indicative of present good health or as predisposing to future good health, the dietary or infant feeding practices that are valued in a given culture, or the health care procedures or facilities for health care that a given authority regards as desirable.

In the latter sense, norms may be used as criteria when evaluating health care, in order to determine the degree of conformity with what is desirable, the average length of stay of patients in hospital, etc. A distinction is sometimes made between norms, defined as quantitative indexes based on research, and standards, which are fixed arbitrarily.

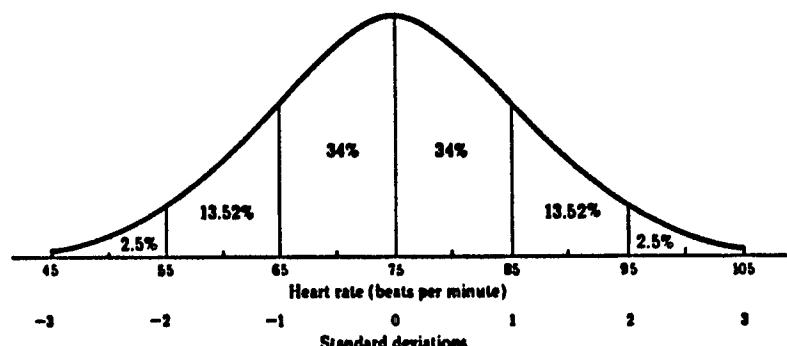
NORMAL This term has three distinct meanings. Conceptual difficulties may arise if these different meanings are not specified, or if the area of their overlap is not clearly understood.

1. Within the usual range of variation in a given population or population group; or frequently occurring in a given population or group. In this sense, "normal" is frequently defined as, "within a range extending from two standard deviations below the mean to two standard deviations above the mean," or "between specified (e.g., the 10th and 90th) percentiles of the distribution."
2. In good health, indicative or predictive of good health, or conducive to good health. For a diagnostic or screening test, a "normal" result is one in a range within which the probability of a specific disease is low (see also NORMAL LIMITS).
3. (Of a distribution) Gaussian; see also NORMAL DISTRIBUTION.

NORMAL DISTRIBUTION (Syn: Gaussian distribution) The continuous frequency distribution of infinite range represented by the equation

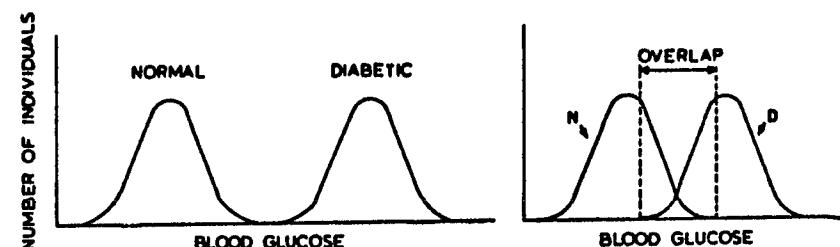
$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(x-\mu)^2/2\sigma^2}$$

where x is the abscissa, $f(x)$ is the ordinate, μ is the mean, σ is the natural logarithm, 2.718 and σ the standard deviation.



The properties of a normal distribution include the following: (1) It is a continuous, symmetrical distribution; both tails extend to infinity; (2) the arithmetic mean, mode, and median are identical; and (3) its shape is completely determined by the mean and standard deviation.

NORMAL LIMITS The limits of the "normal" range of a test or measurement, in the sense of being indicative of or conducive to good health. One way to determine normal limits is to compare the values obtained when the measurements are made in two groups, one that is healthy and has been found to remain healthy, the other ill, or subsequently found to become ill. The result may be two overlapping distributions, as illustrated. Outside the area where the distributions overlap, a given value clearly identifies the presence or absence of disease or some other manifestation of poor health. If a value falls into the area of overlap, the individual may belong to either the normal or the abnormal group. The choice of the normal limits depends upon the relative importance attached to the identification of individuals as healthy or unhealthy. See also FALSE NEGATIVE; FALSE POSITIVE; SENSITIVITY AND SPECIFICITY.



Hypothetical distribution of normal and diabetic glucose levels.
 From Lilienfeld and Lilienfeld, 1979.

NORMATIVE Pertaining to the normal, usual, accepted standard or values. See also NORM.

NOSOCOMIAL Arising while a patient is in a hospital or as a result of being in a hospital; relating to a hospital; denoting a new disorder (unrelated to the patient's primary condition) associated with being in a hospital.

NOSOCOMIAL INFECTION (Syn: hospital-acquired infection) An infection originating in a medical facility, e.g., occurring in a patient in a hospital or other health care facility in whom the infection was not present or incubating at the time of admission. Includes infections acquired in the hospital but appearing after discharge; it also includes such infections among staff.

NOSOGRAPHY, NOSOLOGY Classification of ill persons into groups, whatever the criteria for their classification, and agreement as to the boundaries of the groups, is called "nosology." The assignment of names to each disease entity in the group results in a nomenclature of disease entities, or nosography. (Faber K: *Nosography in Modern Internal Medicine*. New York: Hoeber, 1923.)

NOTIFIABLE DISEASE A disease that, by statutory requirements, must be reported to the public health authority in the pertinent jurisdiction when the diagnosis is made.

A disease deemed of sufficient importance to the public health to require that its occurrence be reported to health authorities.

The reporting to public health authorities of communicable diseases is, unfortunately, very incomplete. The reasons for this include diagnostic inexactitude; the desire of patients and physicians to conceal the occurrence of conditions carrying a social stigma, e.g., sexually transmitted diseases; and the indifference of physicians to the usefulness of information about such diseases as hepatitis, influenza, and measles. Yet notifications are extremely important. They provide the starting point for investigations into the failure of preventive measures such as immunizations, for tracing sources of infection, for finding common vehicles of infection, for describing the geographic clustering of infection, and for various other purposes, depending upon the particular disease.

N.S., n.s. Abbreviation, usually written lower case, for not statistically significant.

NULL HYPOTHESIS (Syn: test hypothesis) The statistical hypothesis that one variable has no association with another variable or set of variables, or that two or more population distributions do not differ from one another. In simplest terms, the null hypothesis states that the results observed in a study, experiment, or test are no different from what might have occurred as a result of the operation of chance alone.

NUMERATOR The upper portion of a fraction used to calculate a rate or a ratio.

NUMERICAL TAXONOMY The construction of homogeneous groupings or taxa using numerical methods; allied to CLUSTER ANALYSIS.

2503012434

OBSERVATIONAL STUDY (Syn: nonexperimental study, survey) Epidemiologic study in situations where nature is allowed to take its course; changes or differences in one characteristic are studied in relation to changes or differences in other(s), without the intervention of the investigator.

OBSERVER VARIATION (ERROR) Variation (or error) due to failure of the observer to measure or to identify a phenomenon accurately. Observer variation erodes scientific credibility whenever it appears. Sir Thomas Browne in *Pseudodoxia Epidemica* (1646), subtitled "Enquiries into very many commonly received tenents and presumed truths," recognized several sources of error: "the common infirmity of human nature, the erroneous disposition of the people, misapprehension, fallacy or false deduction, credulity, obstinate adherence to authority, the belief in popular conceits, the endeavours of Satan."

All observations are subject to variation. Discrepancies between repeated observations by the same observer and between different observers are to be expected; these can be diminished but probably never absolutely eliminated.

Variation may arise from several sources. The observer may miss an abnormality or think he has found one where none is present; a measurement or a test may give incorrect results due to faulty technique or incorrect reading and recording of the results; or the observer may misinterpret the information. Two varieties of observer variation are interobserver variation, i.e., the amount observers vary from one another when reporting on the same material, and intraobserver variation, the amount one observer varies between observations when he reports more than once on the same material.

OCCAM'S RAZOR (Syn: scientific parsimony) William of Occam's 14th century dictum was that, "the assumptions introduced to explain a thing must not be multiplied beyond necessity." This useful maxim does not contradict the conclusion that multiple causes operate in any system. The number of causes implicated depends on the frame of reference of the investigator and on the scope of the inquiry.

OCCURRENCE (Syn: frequency) In epidemiology, a general term describing the frequency of a disease or other attribute or event in a population without distinguishing between INCIDENCE and PREVALENCE.

ODDS The ratio of the probability of occurrence of an event to that of nonoccurrence, or the ratio of the probability that something is so, to the probability that it is not so. If 60 smokers develop a chronic cough and 40 do not, the odds among these 100 smokers in favor of developing a cough are 60:40, or 1.5; this may be contrasted with the probability that these smokers will develop a cough, which is 60/100 or 0.6.

ODDS RATIO (Syn: cross-product ratio, relative odds) The ratio of two odds. The term "odds" is defined differently according to the situation under discussion. Consider

the following notation for the distribution of a binary exposure and a disease in a population or a sample.

| | Exposed | Unexposed |
|------------|---------|-----------|
| Disease | a | b |
| No disease | c | d |

The odds ratio (cross-product ratio) is ad/bc .

The exposure-odds ratio for a set of case control data is the ratio of the odds in favor of exposure among the cases (a/b) to the odds in favor of exposure among noncases (c/d). This reduces to ad/bc . With incident cases, unbiased subject selection, and a "rare" disease (say, under 2% cumulative incidence rate over the study period), ad/bc is an approximate estimate of the RISK RATIO. With incident cases, unbiased subject selection, and DENSITY SAMPLING of controls ad/bc is an estimate of the ratio of the person-time incidence rates (FORCES OF MORBIDITY) in the exposed and unexposed (no rarity assumption is required for this).

The disease-odds (rate-odds) ratio for a cohort or cross section is the ratio of the odds in favor of disease among the exposed (a/c) to the odds in favor of disease among the unexposed (b/d). This reduces to ad/bc and hence is equal to the exposure-odds ratio for the cohort or cross section.

The prevalence-odds ratio refers to an odds ratio derived cross sectionally, as, for example, an odds ratio derived from studies of prevalent (rather than incident) cases.

The risk-odds ratio is the ratio of the odds in favor of getting disease, if exposed, to the odds in favor of getting disease if not exposed. The odds ratio derived from a cohort study is an estimate of this. See also CASE CONTROL STUDY.

ONE-TAIL TEST A statistical significance test based on the assumption that the data have only one possible direction of variability.

OPERATIONAL RESEARCH The systematic study, by observation and experiment, of the working of a system, e.g., health services, with a view to improvement.

OPERATIONS RESEARCH

1. The fitting of models to data, or the designing of models.
2. Synonym for OPERATIONAL RESEARCH.

OPPORTUNISTIC INFECTION Infection with organism(s) that are normally innocuous, e.g., commensals in the human, but become pathogenic when the body's immunologic defenses are compromised, as happens in the acquired immunodeficiency syndrome (AIDS).

ORDINAL SCALE See MEASUREMENT SCALE.

ORDINATE The distance of a point, P , from the horizontal or x axis of a graph, measured along the vertical or y axis. See also ABSCISSA; GRAPH.

OUTCOMES All the possible results that may stem from exposure to a causal factor, or from preventive or therapeutic interventions; all identified changes in health status arising as a consequence of the handling of a health problem. See also CAUSALITY; CAUSATION OF DISEASE, FACTORS IN.

OUTLIERS Observations differing so widely from the rest of the data as to lead one to suspect that a gross error may have been committed, or suggesting that these values come from a different population.

OUTBREAK (Syn: epidemic) Sometimes the preferred word, as it may escape sensationalism associated with the word epidemic. Alternatively, a localized as opposed to generalized epidemic.

OUTPUT The immediate result of professional or institutional health care activities, usu-

ally expressed as units of service, e.g., patient hospital days, outpatient visits, laboratory tests performed.

OVERMATCHING A situation that may arise when groups are matched. Several varieties can be distinguished:

1. The matching procedure partially or completely obscures evidence of a true causal association between the independent and dependent variables. Overmatching may occur if the matching variable is involved in, or is closely connected with, the mechanism whereby the independent variable affects the dependent variable. The matching variable may be an intermediate cause in the causal chain or it may be strongly affected by, or a consequence of, such an intermediate cause.
2. The matching procedure uses one or more unnecessary matching variables, e.g., variables that have no causal effect or influence on the dependent variable, and hence cannot confound the relationship between the independent and dependent variables.
3. The matching process is unduly elaborate, involving the use of numerous matching variables and/or insisting on very close similarity with respect to specific matching variables. This leads to difficulty in finding suitable controls. See also MATCHING.

OVERWINTERING See VECTOR-BORNE INFECTION.

P

P, P (PROBABILITY) VALUE The probability that a test statistic would be as extreme as or more extreme than observed if the null hypothesis were true. The letter P, followed by the abbreviation n.s. (not significant) or by the symbol < (less than) and a decimal notation such as 0.01, 0.05, is a statement of the probability that the difference observed could have occurred by chance, if the groups are really alike, i.e., under the NULL HYPOTHESIS.

Investigators may arbitrarily set their own significance levels, but in most biomedical and epidemiologic work, a study result whose probability value is less than 5% ($P < 0.05$) or 1% ($P < 0.01$) is considered sufficiently unlikely to have occurred by chance to justify the designation "statistically significant." See also STATISTICAL SIGNIFICANCE.

PAIRED SAMPLES In a CLINICAL TRIAL, pairs of subject patients may be studied. One member of each pair receives the experimental regimen, and the other receives a suitably designated control regimen. Pairing should be based on a prognostic variable such as age.

Pairing may similarly be used in a CASE CONTROL STUDY or in a COHORT STUDY. See also MATCHING.

PANDEMIC An epidemic occurring over a very wide area and usually affecting a large proportion of the population.

PANEL STUDY A combination of cross-sectional and cohort methods, in which the investigator conducts a series of cross-sectional studies of the same individuals or study sample. This method of study permits changes in one variable to be related to changes in other variables. See also NESTED CASE-CONTROL STUDY.

PANUM, PETER LUDWIG (1820-1885) A Danish physician who observed firsthand an epidemic of measles in the Faroe Islands in 1846. This was the first outbreak there for many years, and from the epidemic pattern, Panum deduced some basic, previously unknown details about the method of spread, and incubation period, the lasting immunity that followed infection, and the relationship between age and severity of infection.

PARADIGM A typical example, a pattern of thought or conceptualization; an overall way of regarding phenomena, within which scientists normally work. A paradigm may dictate what form of explanation will be found acceptable, but a science may change paradigms. In many contexts in which it is used, the term is both ambiguous and vague.¹ The word is often used loosely as a synonym for "factor" or "variable."

¹Kuhn T. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

PARAMETER In mathematics, a constant in a formula or model; in statistics and epidemiology, a measurable characteristic of a population.

PARAMETRIC TEST A statistical test that depends upon assumption(s) about the distribution of the data, e.g., that these are normally distributed.

PARASITE An animal or vegetable organism that lives on or in another and derives its nourishment therefrom. An obligate parasite is one that cannot lead an independent nonparasitic existence. A facultative parasite is one that is capable of either parasitic or independent existence.

PARASITE COUNT See WORM COUNT.

PARASITE DENSITY The collective degree of parasitemia in a population, calculated by the use of either the geometric mean or the weighted average of the individual parasite counts; e.g., by using a frequency distribution based on a geometric progression.

PARATENIC HOST (Syn: transport host) A second, third, or subsequent intermediate host of a parasite, in which the parasite does not undergo any development or replication, but remains, usually encysted, until the paratenic host is ingested by the definitive host of the parasite.

PARITY The status of a woman as regards the fact of having borne viable children. The number of full-term children previously borne by a woman, excluding miscarriages or abortions in early pregnancy, but including stillbirths.

PARTICULARIZATION A method of analysis opposite to generalization or abstraction. It focuses on the specificity of a number of facts and illustrates an issue through the use of example.

PASSAGE The transfer of micro-organisms from human to animal host(s) either directly or via laboratory culture; in the laboratory, this procedure is used to establish the Henle-Koch postulates.

PASSENGER VARIABLE A variable that varies systematically with the dependent variable under study, without being causally related to it; a third (explanatory) variable, the common cause of both the dependent and the passenger variable, "explains" or accounts for their association.

PASSIVE SMOKING See INVOLUNTARY SMOKING.

PASTEUR, LOUIS (1822-1895) A French chemist and biologist. One of the founders of bacteriology and therefore an important figure also in epidemiology. Starting in chemistry, he worked out the biological basis for fermentation, and then went on to make many important discoveries in bacteriology, notably vaccines against anthrax and rabies. He is, of course, eponymously honored by the word "pasteurization."

PATH ANALYSIS A mode of analysis involving assumptions about the direction of causal relationships between linked sequences and configurations of variables. This permits the analyst to construct and test the appropriateness of alternative models (in the form of a path diagram) of the causal relations that may exist within the array of variables included in the finite system studied. Identification of the less probable sequences of causal pathways may permit them to be eliminated from further consideration.

PATHOGEN Organism capable of causing disease (literally, causing a pathological process).

PATHOGENESIS The postulated mechanisms by which the etiologic agent produces disease. The difference between ETIOLOGY and pathogenesis should be noted: The etiology of a disease or disability consists of the postulated causes that initiate the pathogenetic mechanisms; control of these causes might lead to prevention of the disease.

PATHOGENICITY The property of an organism that determines the extent to which overt disease is produced in an infected population, or the power of an organism to produce disease. Also used to describe comparable properties of toxic chemicals,

etc. Pathogenicity of infectious agents is measured by the ratio of the number of persons developing clinical illness to the number exposed to infection. See also **VIRULENCE**, with which pathogenicity is sometimes confused.

PEARSON, KARL (1857-1936) British mathematician, biologist and geneticist. Pearson was a pupil of Francis Galton, who led the science of statistics further into applications in biology and genetics. He founded the journal *Biometrika*, coined the word "biometry," and taught the next generation of statistician/epidemiologists, including Major Greenwood, Raymond Pearl, and others.

PEARSON'S PRODUCT MOMENT CORRELATION See **CORRELATION COEFFICIENT**.

PEDIGREE A diagram showing the ancestral relationships and transmission of genetic traits over several generations of a family.

PEER REVIEW Process of review of research proposals, manuscripts submitted for publication, abstracts submitted for presentation at scientific meetings, whereby these are judged for scientific and technical merit by other scientists in the same field. Also refers to review of clinical performance, when it is a form of medical audit.

PENETRANCE The frequency, expressed as a percentage, with which individuals of a given phenotype manifest at least some degree of a specific mutant phenotype associated with a trait. See also **GENETIC PENETRANCE**.

PERCEIVED NEED A felt need. The term usually refers to need for health care that is felt by the person or community concerned, but which may not be perceived by health professionals.

PERCENTILE The set of divisions that produce exactly 100 equal parts in a series of continuous values, such as children's heights or weights. Thus a child above the 90th percentile has a greater value for height or weight than over 90% of all in the series.

PERINATAL MORTALITY Literally, mortality around the time of birth. Conventionally this time is limited to the period between 28 weeks gestation and one week postnatal. However, as the following discussion indicates, other factors, especially the weight of the fetus, should be considered. The *Ninth (1975) Revision of the International Classification of Diseases* includes the following:

Perinatal mortality statistics

It is recommended that national perinatal statistics should include all fetuses and infants delivered weighing at least 500 g (or, when birth weight is unavailable, the corresponding gestational age [22 weeks] or body length [25 cm crown-heel]), whether alive or dead. It is recognized that legal requirements in many countries may set different criteria for registration purposes, but it is hoped that countries will arrange the registration or reporting procedures in such a way that the events required for inclusion in the statistics can be identified easily. It is further recommended that less mature fetuses and infants should be excluded from perinatal statistics unless there are legal or other valid reasons to the contrary.

It is recommended above that national statistics would include fetuses and infants weighing between 500 g and 1000 g both for their inherent value and because their inclusion improves the completeness of reporting at 1000 g and over.

Inclusion of this group of very immature births, however, disrupts international comparisons because of differences in national practices concerning their registration. Another factor affecting international comparisons is that all live-born infants, irrespective of birth weight, are included in the calculation of rates, whereas some lower limit of maturity is applied to infants born dead.

In order to eliminate these factors, it is recommended that countries should pres-

ent, solely for international comparisons, "standard perinatal statistics" in which both the numerator and denominator of all rates are restricted to fetuses and infants weighing 1000 g or more (or, where birth weight is unavailable, the corresponding gestational age [28 weeks] or body length [25 cm crown-heel]).

PERINATAL MORTALITY RATE In most industrially developed nations, this is defined as

$$\text{Perinatal mortality rate} = \frac{\text{Fetal deaths (28 weeks + of gestation) + postnatal deaths (first week)}}{\text{Fetal deaths (28 weeks + of gestation) + live births}} \times 1000$$

The World Health Organization's definition, more appropriate in nations with less well-established vital records, is

$$\text{Perinatal mortality rate} = \frac{\text{Late fetal deaths (28 weeks + of gestation) + postnatal deaths (first week)}}{\text{Live births in a year}} \times 1000$$

Note the differences in denominator of the perinatal mortality rate as defined by WHO and in industrially developed nations. This makes international comparison difficult. The WHO Expert Committee on the Prevention of Perinatal Mortality and Morbidity (1970) recommended a more precise formulation: "Late fetal and early neonatal deaths weighing over 1000 g at birth expressed as a ratio per 1000 live births weighing over 1000 g at birth."

PERIODIC (MEDICAL) EXAMINATIONS Assessment of health status conducted at predetermined intervals, e.g., annually or at specified milestones in life such as infancy, school entry, preemployment, or preretirement. This form of medical examination generally follows a formal protocol, e.g., employing a set of structured questions and/or a predetermined set of laboratory tests.

PERIOD OF COMMUNICABILITY See **COMMUNICABLE PERIOD**.

PERMISSIBLE EXPOSURE LIMIT (PEL) An occupational health standard to safeguard employees against dangerous chemicals or contaminants in the workplace. See **SAFETY STANDARDS**.

PERSONAL HEALTH CARE Those services to individuals that are performed on a one-to-one basis by a health care worker for the purpose of maintaining or restoring health.

PERSONAL MONITORING DEVICE An instrument attached to a person to measure the exposure of that person to hazardous substance(s).

PERSON-TIME A measurement combining persons and time, used as denominator in person-time incidence and mortality rates. It is the sum of individual units of time that the persons in the study population have been exposed to the condition of interest. A variant is person-distance, e.g., as in passenger-kilometers. The most frequently used person-time is person-years. With this approach, each subject contributes only as many years of observation to the population at risk as he is actually observed; if he leaves after one year, he contributes one person-year; if after ten, ten person-years. The method can be used to measure incidence over extended and variable time periods.

PERSON-TIME INCIDENCE RATE (Syn: interval incidence density) A measure of the incidence rate of an event, e.g., a disease or death, in a population at risk, given by

person-to-person spread

98

$$\frac{\text{Number of events occurring during the interval}}{\text{Number of person-time units at risk observed during the interval}}$$

PERSON-TO-PERSON SPREAD OF DISEASE (Syn: prosodemic) See **TRANSMISSION OF INFECTION**.

PERSON-YEARS See **PERSON-TIME**.

PETTY, WILLIAM (1623–1687) A member of the same circle as John Graunt, he is equally recognized as a pioneer in vital statistics and economics. His ideas and concepts of lifetime earning capability are contained in *Political Arithmetic* (London, 1691).

PHARMACOEPIDEMIOLOGY The study of the distribution and determinants of drug-related events in populations, and the application of this study to efficacious drug treatment.

PHYSICIAN (Syn: medical practitioner, doctor) Professional person qualified by education and authorized by law to practice medicine.

PIE CHART A circular diagram divided into segments, each representing a category or subset of data. The amount for each category is proportional to the angle subtended at the center of the circle and hence to the area of the sector.

When several pie charts are used to describe several populations, the area of each circle is proportional to the size of the population it represents.

PILOT INVESTIGATION, STUDY A small-scale test of the methods and procedures to be used on a larger scale if the pilot study demonstrates that these methods and procedures can work.

PLACEBO, PLACEBO EFFECT An inert medication or procedure. The placebo effect (usually but not necessarily beneficial) is attributable to the expectation that the regimen will have an effect, i.e., the effect is due to the power of suggestion. See also **HALO EFFECT**.

POINT SOURCE EPIDEMIC See **EPIDEMIC, COMMON SOURCE**.

POISSON DISTRIBUTION A distribution function used to describe the occurrence of rare events or to describe the sampling distribution of isolated counts in a continuum of time or space (e.g., sample counts of radioactive disintegration per minute). The number of events has a Poisson distribution with parameter λ (lambda) if the probability of observing k events ($k = 0, 1, \dots$) is equal to

$$p(x=k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

where e is the base of natural logarithm, 2.7183. . . . The mean and variance of the distribution are both equal to λ . This distribution is used in modeling person-time incidence rates.

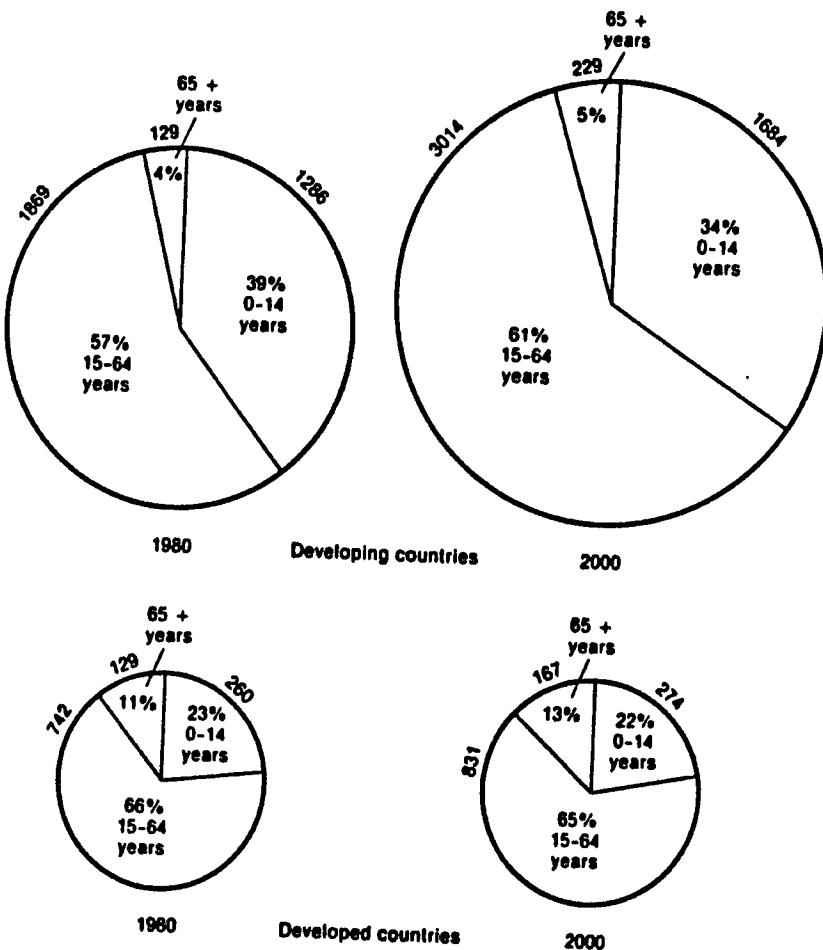
POLLUTION Any undesirable modification of air, water, or food by substance(s) that are toxic or may have adverse effects on health or that are offensive though not necessarily harmful to health.

POLYGENIC INHERITANCE The transmission of a phenotypic trait whose expression depends upon the additive effect of a number of genes.

POUNDERAL INDEX The anthropometric index of body mass. Defined as height divided by the cube root of the body weight. The **BODY MASS INDEX** is generally regarded as a better index of body mass.

99

population attributable risk



Pie charts of age structure of the population.
(Figures outside the circle show the population in millions.)
From World Health Organization.

POPULATION

1. All the inhabitants of a given country or area considered together; the number of inhabitants of a given country or area.
2. (In sampling) The whole collection of units from which a sample may be drawn; not necessarily a population of persons; the units may be institutions, records, or events. The sample is intended to give results that are representative of the whole population.

POPULATION ATTRIBUTABLE RISK (PAR) This term is used by many epidemiologists^{1,2,3} in preference to the terms "attributable fraction (population)" or "etiological fraction"

population attributable risk percent 100

(population)." It is the incidence of a disease in a population that is associated with (attributable to) exposure to the risk factor. It is often expressed as a percentage. It is calculated by similar methods to those described for attributable fraction (population), i.e.,

$$\text{PAR\%} = \frac{P_e(I_e - I_u)}{P_e \times I_t} \times 100$$

where P_e = number of persons exposed

P_e = persons in the population

I_e = incidence rate among the exposed

I_u = incidence rate among the unexposed

I_t = incidence rate for the total population

In a case-control study, PAR can be estimated in various ways; Cole and MacMahon³ give the following formula:

$$\text{PAR\%} = \frac{P_e(RR - 1)}{1 + P_e(RR - 1)} \times 100$$

where P_e = proportion of controls exposed

RR = relative risk for exposed, compared to risk of 1 for the unexposed.

¹MacMahon B, Pugh TF: *Epidemiology: Principles and Methods*. Boston: Little, Brown, 1970.

²Fletcher RH, Fletcher SW, Wagner EH: *Clinical Epidemiology—the Essentials*. Baltimore: Williams & Wilkins, 1982.

³Cole P, MacMahon B: Attributable risk percent in case-control studies. *Br J Prev Soc Med* 25:242-244, 1971.

POPULATION ATTRIBUTABLE RISK PERCENT This is the attributable fraction in the population, expressed as a percentage. See also ATTRIBUTABLE FRACTION (POPULATION).

POPULATION BASED Pertaining to a general population defined by geopolitical boundaries; this population is the denominator and/or the sampling frame.

POPULATION DYNAMICS Changes in the structure of a population; loosely used as a synonym for demography.

POPULATION EXCESS RATE A measure of the amount of disease associated with exposure to a putative cause of the disease in the population. It is the difference between the rates of disease in the entire population and among the nonexposed.

POPULATION MEDICINE See COMMUNITY MEDICINE.

POPULATION MOMENTUM In a growing population, the phenomenon of continuing population growth beyond the time when replacement level fertility has been achieved, because of the increasing size of child-bearing and younger age cohorts, resulting from higher fertility and/or falling mortality in preceding years.

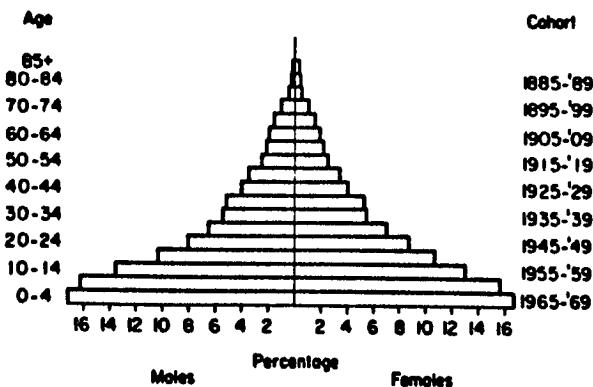
POPULATION PYRAMID A graphic presentation of the age and sex composition of the population. The population pyramid is constructed by computing the percentage distribution of a population, simultaneously cross-classified by sex and age. The percentage that each female age group is of the total is plotted on the right and the corresponding percentages for males are plotted on the left. A population pyramid is intended to provide a quick overall comprehension of age and sex structure in the population. A population whose pyramid has a broad base and narrow apex may be identified as a high fertility population. Changing shape over time reflects the changing composition of the population, associated with changes in fertility and mortality at each age.

Since the figure is two dimensional, the word "pyramid" is incorrectly used, but the more accurate word "profile" has never caught on.

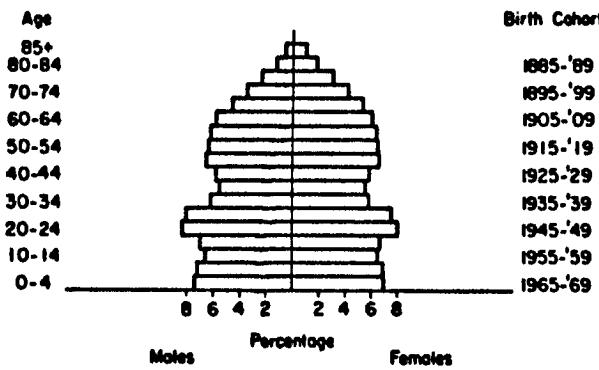
101

postneonatal mortality rate

Mexico 1970



Sweden 1970



Population pyramids.

Top: High fertility, low proportion survive to old age (Mexico).

Bottom: Low fertility, high proportion survive to old age (Sweden).

From Last, 1980.

POPULATION, STUDY The group selected for investigation.

POPULATION, TARGET The group from which a study population is selected.

POSTERIOR ODDS, POSTERIOR PROBABILITY Probability calculated after reference to results of a study. See BAYES THEOREM.

POSTMARKETING SURVEILLANCE A procedure implemented after a drug has been licensed for public use, designed to provide information on the actual use of the drug for a given indication, and on the occurrence of side effects, adverse reactions, etc. A method for epidemiologic study of adverse drug reactions.

POSTNEONATAL MORTALITY RATE The number of infant deaths between 28 days and one year of age in a given year per 1000 live births in that year. It is an important

rate to monitor in developing countries where older infants frequently die of infections and malnutrition.

POTENCY The strength of a particular drug, toxin, or hazard; the ratio of the dose of a standard amount required to elicit a specific response, to the dose of the test agent that elicits the same response.

POTENTIAL YEARS OF LIFE LOST (PYLL) A measure of the relative impact of various diseases and lethal forces on society. PYLL highlights the loss to society as a result of youthful or early deaths. The figure for potential years of life lost due to a particular cause is the sum, over all persons dying from that cause, of the years that these persons would have lived had they experienced normal life expectation. The concept derives from Petty's *Political Arithmetic* (1687) and is elaborated upon in Dublin and Lotka's *Money Value of a Man* (1930).

POWER A characteristic of a statistical hypothesis test, denoting the probability that the null hypothesis will be rejected if it is indeed false. It is equal to 1 minus the probability of type II error. See also **ERROR**, **RESOLUTION**. Resolving power is the comparable property of individual measurements.

PRAGMATIC STUDY A study whose aim is to improve health status or health care of a specified population, provide a basis for decisions about health care, or evaluate previous actions. See also **EXPLANATORY STUDY**; **COMMUNITY DIAGNOSIS**; **PROGRAM REVIEW**.

PRECISION

1. The quality of being sharply defined or stated. One measure of precision is the number of distinguishable alternatives from which a measurement was selected, sometimes indicated by the number of significant digits in the measurement. Another measure of precision is the standard error of measurement, the standard deviation of a series of replicate determinations of the same quantity. Precision does not imply accuracy. See also **MEASUREMENT**, **PROBLEMS WITH TERMINOLOGY**.
2. In statistics, precision is defined as the inverse of the variance of a measurement or estimate.

PRECURSOR An early stage in the course of a disease, or a condition or state preceding pathological onset of a disease; sometimes detectable by **SCREENING**; may be identified as a **RISK MARKER**.

PREDICTIVE VALUE In screening and diagnostic tests, the probability that a person with a positive test is a true positive (i.e., does have the disease) is referred to as the "predictive value of a positive test." The predictive value of a negative test is the probability that a person with a negative test does not have the disease. The predictive value of a screening test is determined by the sensitivity and specificity of the test, and by the prevalence of the condition for which the test is used. See also **SCREENING**; **SENSITIVITY AND SPECIFICITY**.

PREMUNITION A term used mainly in the epidemiology of parasitic diseases, especially malaria. It signifies a state of resistance, in a host harboring a parasite, to superinfection by a parasite of the same species. This state is dependent on the continued survival of parasites in the body and disappears after their elimination. It may be complete or partial.

PREPATENT PERIOD In parasitology, the period equivalent to the incubation period of microbial infections; the corresponding phase may be biologically different from microbial multiplication when the invading organism is a multicellular parasite that undergoes developmental stages in the host.

PRESCRIPTIVE SCREENING See **SCREENING**.

PREVALENCE The number of instances of a given disease or other condition in a given population at a designated time; sometimes used to mean **PREVALENCE RATE**. When used without qualification the term usually refers to the situation at a specified point in time (point prevalence).

prevalence, annual (An occasionally used index) The total number of persons with the disease or attribute at any time during a year. It includes cases of the disease arising before but extending into or through the year as well as those having their inception during the year.

prevalence, lifetime The total number of persons known to have had the disease or attribute for at least part of their life.

prevalence, period The total number of persons known to have had the disease or attribute at any time during a specified period.

prevalence, point The number of persons with a disease or an attribute at a specified point in time.

PREVALENCE RATE (RATIO) The total number of all individuals who have an attribute or disease at a particular time (or during a particular period) divided by the population at risk of having the attribute or disease at this point in time or midway through the period. A problem may arise with calculating period prevalence rates because of the difficulty of defining the most appropriate denominator. See also **PREVALENCE**.

PREVALENCE STUDY See **CROSS-SECTIONAL STUDY**.

PREVENTABLE FRACTION (population) In a situation in which exposure to a given factor is believed to protect against a disease (or other outcome), the preventable fraction in the population is the proportion of the disease (in the population) that would be prevented if the whole population were exposed to the factor. This value must be interpreted with caution, as part or all of the apparent protective effect may be due to other factors associated with the apparent protective factor.

In a study of a total population, the preventable fraction (population) is computed as $\frac{I_p - I_e}{I_p}$, where I_p is the incidence rate of the disease (or other outcome)

in the population, and I_e is the incidence rate in the exposed persons in the population.

PREVENTED FRACTION (population) In a situation in which exposure to a given factor is believed to protect against a disease (or other outcome), the prevented fraction is the proportion of the hypothetical total load of disease (in the population) that has been prevented by exposure to the factor. This value must be interpreted with caution, as part or all of the apparent protective effect may be due to other factors associated with the apparent protective factor.

In a study of a total population the prevented fraction is computed as $\frac{I_u - I_p}{I_u}$,

where I_p is the rate of the disease in the population, and I_u is the rate among people unexposed to the factor.

PREVENTION The goals of medicine are to promote health, to preserve health, to restore health when it is impaired, and to minimize suffering and distress. These goals are embodied in the word "prevention," which is easiest to define in the context of levels, customarily called primary, secondary, and tertiary prevention. Authorities on **PREVENTIVE MEDICINE** do not agree on the precise boundaries between these levels, nor on how many levels can be distinguished, but the differences of opinion are semantic rather than substantive.

An epidemiologic interpretation of the distinction between primary and second-

ary prevention is that primary prevention is aimed at reducing incidence of disease and other departures from good health, secondary prevention aims to reduce prevalence by shortening the duration, and tertiary prevention is aimed at reducing complications.

Primary prevention can be defined as the protection of health by personal and community-wide effects, e.g., preserving good nutritional status, physical fitness, and emotional well-being, immunizing against infectious diseases, and making the environment safe. (But see also **HEALTH PROMOTION**.)

Secondary prevention can be defined as the measures available to individuals and populations for the early detection and prompt and effective intervention to correct departures from good health.

Tertiary prevention consists of the measures available to reduce or eliminate long-term impairments and disabilities, minimize suffering caused by existing departures from good health, and to promote the patient's adjustment to irremediable conditions. This extends the concept of prevention into the field of rehabilitation.

PREVENTIVE MEDICINE The application of preventive measures by clinical practitioners. A specialized field of medical practice composed of distinct disciplines that utilize skills focusing on the health of defined populations in order to promote and maintain health and well-being and prevent disease, disability, and premature death.

In addition to the knowledge of basic and clinical sciences and the skills common to all physicians, the distinctive aspects of preventive medicine include knowledge of and competence in biostatistics, epidemiology, administration including planning, organization, management, financing, and evaluation of health programs; environmental health; application of social and behavioral factors in health and disease; and the application of primary, secondary, and tertiary prevention measures within clinical medicine. (The above is the definition and description of the field that has been adopted by the American College of Preventive Medicine; for completeness, at least two other items ought to be added, i.e., health education and nutrition).

PRIMARY CASE The individual who introduces the disease into the family or group under study. Not necessarily the first diagnosed case in a family or group. See also **INDEX CASE**.

PRIMARY HEALTH CARE

1. Health care that begins at the time of first encounter between a patient and a provider of health care; An alternative term is primary medical care.
2. The WHO definition of primary health care includes much more: Primary health care is essential health care made accessible at a cost the country and the community can afford, with methods that are practical, scientifically sound, and socially acceptable. Everyone in the community should have access to it, and everyone should be involved in it. Related sectors should also be involved in it in addition to the health sector. At the very least it should include education of the community on the health problems prevalent and on methods of preventing health problems from arising or of controlling them; the promotion of adequate supplies of food and of proper nutrition; sufficient safe water and basic sanitation; maternal and child health care including family planning; the prevention and control of locally endemic diseases; immunization against the main infectious diseases; appropriate treatment of common diseases and injuries; and the provision of essential drugs. (From *Glossary of Terms Used in the Health for All Series No. 1-8*. Geneva: WHO, 1984.)

PRINCIPAL COMPONENT ANALYSIS A statistical method to simplify the description of a

set of interrelated variables. Its general objectives are data reduction and interpretation; there is no separation into dependent and independent variables; the original set of correlated variables is transformed into a smaller set of uncorrelated variables called the principal components. Often used as the first step in a factor analysis.

PRIOR PROBABILITY Probability calculated or estimated from theory or belief, before a study is done. See **BAYES' THEOREM**.

PROBABILITY

1. The limit of the relative frequency of an event in a sequence of N random trials as N approaches infinity, i.e., the limit of

$$\frac{\text{Number of occurrence of the event}}{N}$$

2. A measure, ranging from zero to 1, of the degree of belief in a hypothesis or statement.

PROBABILITY DENSITY The frequency distribution of a continuous random variable.

PROBABILITY DISTRIBUTION For a discrete random variable, the function that gives the probabilities that the variable equals each of a sequence of possible values. Examples include the binomial and Poisson distributions. For a continuous random variable, often used synonymously with the probability density function.

PROBABILITY SAMPLE (Syn: random sample) See **SAMPLE**.

PROBABILITY THEORY The branch of mathematics dealing with the purely logical properties of probability. Its theorems underly most statistical methods.

PROBAND See **PROPOSITUS**.

PROBLEM-ORIENTED MEDICAL RECORD (POMR) A medical record in which the patient's history, physical findings, laboratory results, etc., are organized to give a cumulative record of problems, e.g., hemoptysis, rather than disease, e.g., pneumonia. The record includes subjective, objective, and significant negative information, discussions and conclusions, and diagnostic and treatment plans with respect to each problem. The record, which was developed by Lawrence Weed,¹ contrasts with the traditional medical record, which is less formally organized, usually recording all information from each source (history, physical, and laboratory findings) together without regard to the problems the information describes.

Since the problems may not be described in terms of conventional disease labels, their classification and counting for epidemiologic purposes are sometimes difficult. The **INTERNATIONAL CLASSIFICATION OF HEALTH PROBLEMS IN PRIMARY CARE (ICHPPC)** is an attempt to overcome this difficulty.

¹Weed LL: Medical records that guide and teach. *New Engl J Med* 278:593-600, 652-657, 1968.

PROCATARCTIC CAUSE A term used by epidemiologists of the late 19th and early 20th centuries, probably last used by GREENWOOD, to describe predisposing causes associated with habits of life.

PROFESSIONAL ACTIVITY STUDY (PAS) The **HOSPITAL DISCHARGE ABSTRACT SYSTEM** that covers many acute short-stay hospitals in the United States. It provides regularly published statistical tables arranged according to hospital service, diagnostic category, etc., giving details on diagnostic and therapeutic procedures, length of stay and outcome.

PROGRAM

1. A (formal) set of procedures to conduct an activity, e.g., control of malaria.
2. An ordered list of instructions directing a computer to carry out a desired sequence of operations. The objective is normally the solution of a problem.

PROGRAM EVALUATION AND REVIEW TECHNIQUES (PERT) A work-scheduling method that uses ALGORITHMS and also enunciates general principles of procedure for allocating resources. Calls for listing specific tasks to be completed and the resources—personnel, equipment, supplies, and other items—that will be needed, along with their costs, a time chart indicating when each component task is to begin and end, giving interim accomplishment levels during that period, and a specification of times for interim review of the progress of the plan.

PROGRAM REVIEW An evaluative study of a specific health program operating in a specific setting, performed to provide a basis for decisions concerning the operation of the program.

PROGRAM TRIAL An experimental or quasi-experimental evaluative study of a (health) program.

PROLECTIVE Pertaining to data collected by planning in advance. Contrast retrolective. The terms prolective and retrolective, coined by AR Feinstein¹ are said to describe more precisely the actions of research workers than the common terms prospective and retrospective; use of these terms is limited, and is deprecated by many epidemiologists.

¹Clin Pharmacol Ther 30:564-577, 1981.

PROPORTION A type of ratio in which the numerator is included in the denominator. The ratio of a part to the whole, expressed as a "decimal fraction" (e.g., 0.2), as a "vulgar fraction" ($\frac{1}{5}$), or, loosely, as a percentage (20%). By definition, a proportion (p) must be in the range (decimal) $0.0 \leq p \leq 1.0$. Since numerator and denominator have the same dimension, any dimensional contents cancel out, and a proportion is a dimensionless quantity. Where numerator and denominator are based upon counts rather than upon measurements, the originals are also dimensionless, although it should be understood that proportions can be used for measured quantities, e.g., the skin area of the lower limb is x percent of the total skin area, as well as for counts, e.g., 0.15 of the population died. A prevalence rate is a count-based proportion. The nondimensionality of a proportion, and its range limitations, do not necessarily apply to other kinds of ratios, of which "proportion" is a subset. See also RATE; RATIO.

PROPORTIONAL HAZARDS MODEL (Syn: Cox model) A statistical MODEL IN SURVIVAL ANALYSIS that asserts that the effect of the study factors on the HAZARD RATE in the study population is multiplicative and does not change over time. For example, the model for two factors x_1 and x_2 asserts that the rate at time t $\lambda(t)$, is given by

$$e^{\beta_{1x_1} + \beta_{2x_2}} \lambda_0(t)$$

where $\lambda_0(t)$ is the rate when $x_1 = x_2 = 0$, and e is the (natural) exponential function.

PROPORTIONATE MORTALITY RATE, RATIO (PMR) Number of deaths from a given cause in a specified time period, per 100 or 1000 total deaths in the same time period. Can give rise to misleading conclusions if used to compare mortality experience of populations with different distributions of causes of death.

PROPOSITUS (Syn: proband) The family member who first draws attention to a (genetic) pedigree of a given trait. The INDEX CASE in a genetic study.

PROSPECTIVE STUDY See COHORT STUDY.

PROTOCOL The plan, or set of steps, to be followed in a study or investigation, or in an intervention program. See also ALGORITHM, CLINICAL.

PROXIMATE DETERMINANT OF FERTILITY Factor having a direct influence on fertility; such factors include age at marriage, breastfeeding, abortion, and contraceptive use.

PUBLIC HEALTH Public health is one of the efforts organized by society to protect, promote, and restore the people's health. It is the combination of sciences, skills, and beliefs that is directed to the maintenance and improvement of the health of all the people through collective or social actions. The programs, services, and institutions involved emphasize the prevention of disease and the health needs of the population as a whole. Public health activities change with changing technology and social values, but the goals remain the same: to reduce the amount of disease, premature death, and disease-produced discomfort and disability in the population. Public health is thus a social institution, a discipline, and a practice.

PUNCH CARD A card on which data are stored by means of holes punched in specified positions; useful in storing, processing, and analyzing data. Edge-punch cards have marginal holes converted to slots by punching so that they can be manually sorted. The commonly used variety of punch cards have 80 columns and 12 rows. In each column of the card there are 12 positions at which holes may be punched, according to a predetermined code. The position of the hole is the means of identifying the value of a variable. Punch cards of this type are sorted mechanically or electrically to provide a rapid means of processing and analyzing data, sometimes of great complexity. See also DATA PROCESSING.

P VALUE See P (PROBABILITY).

QALY Acronym for quality-adjusted life years; this is an adjustment of life expectancy that allows for prevalence of activity-limitation, assessed from hospital discharge data or by health survey data, in the population subgroup for which QALY is calculated. For example, the life expectancy of males at birth in Canada in 1978 was 70.8 years; after adjusting for activity-limitation using health survey data, quality-adjusted life expectancy, or QALY, was 65.8 years.¹

¹Wilkins R, Adams O: *Healthfulness of Life*. Montreal, 1983.

QUALITATIVE DATA Observations or information characterized by measurement on a categorical scale, i.e., a dichotomous or nominal scale, or, if the categories are ordered, an ordinal scale. Examples are sex, hair color, death or survival, and nationality. See also MEASUREMENT SCALE.

QUALITY CONTROL The supervision and control of all operations involved in a process, usually involving sampling and inspection, in order to detect and correct systematic or excessively random variations in quality.

QUALITY OF CARE A level of performance or accomplishment that characterizes the health care provided. Ultimately, measures of the quality of care always depend upon value judgments, but there are ingredients and determinants of quality that can be measured objectively. These ingredients and determinants have been classified by Donabedian¹ into measures of structure (e.g., manpower, facilities), process (e.g., diagnostic and therapeutic procedures), and outcome (e.g., case fatality rates, disability rates, and levels of patient satisfaction with the service). See also HEALTH SERVICES RESEARCH.

¹Donabedian A: *A Guide to Medical Care Administration* (Vol. 2). New York: American Public Health Association, 1969.

QUALITY OF LIFE In a general sense, that which makes life worth living. In a more "quantitative" sense, an estimate of remaining life free of impairment, disability or handicap, as used in the expression "quality adjusted life years;" somewhere between these is an estimate of the utility of life—for instance, in clinical decision analysis, the utility of life that is impaired by a disabling degree of angina pectoris may be compared with that of a life that may be shorter in duration but free of disabling pain, as a result of applying therapeutic procedures. Such trade-offs are part of clinical decision analysis. See also UTILITY.

QUANTILES Divisions of a distribution into equal, ordered subgroups. Deciles are tenths; quartiles, quarters; quintiles, fifths; terciles, thirds; and centiles, hundredths.

QUANTITATIVE DATA Data in numerical quantities such as continuous measurements or counts.

QUARANTINE The 14th edition of *Control of Communicable Disease in Man*¹ gives the following:

Restriction of the activities of well persons or animals who have been exposed to

a case of communicable disease during its period of communicability (i.e., contact) to prevent disease transmission during the incubation period if infection should occur.

- a) **Absolute or complete quarantine:** The limitation of freedom of movement of those exposed to a communicable disease for a period of time not longer than the longest usual incubation period of that disease, in such manner as to prevent effective contact with those not so exposed (see Isolation).
- b) **Modified quarantine:** A selective, partial limitation of freedom of movement of contacts, commonly on the basis of known or presumed differences in susceptibility and related to the danger of disease transmission. It may be designed to meet particular situations. Examples are exclusion of children from school, exemption of immune persons from provisions applicable to susceptible persons, or restriction of military populations to the post or to quarters. It includes: Personal surveillance, the practice of close medical or other supervision of contacts in order to permit prompt recognition of infection or illness but without restricting their movements; and Segregation, the separation of some part of a group of persons or domestic animals from the others for special consideration, control or observation—removal of susceptible children to homes of immune persons, or establishment of a sanitary boundary to protect uninfected from infected portions of a population.

See also ISOLATION.

¹Washington DC: American Public Health Association, 1985.

QUASI-EXPERIMENT An experiment in which the investigator lacks full control over the allocation and/or the timing of the intervention.

QUESTIONNAIRE A predetermined set of questions used to collect data—clinical data, social status, occupational group, etc. This term is often applied to a self-completed survey instrument, as contrasted with an INTERVIEW SCHEDULE.

QUETELET, LAMBERT ADOLPHE JACQUES (1796–1857) Belgian astronomer, statistician, and social scientist, one of the first to apply statistical thinking to the social and biological sciences, e.g., in delineating the (normal) distribution of variables such as height in the population. He influenced others who followed, e.g., FLORENCE NIGHTINGALE.

QUETELET'S INDEX See BODY MASS INDEX.

QUOTA SAMPLING A method by which the proportions in the sample in various subgroups (according to criteria such as age, sex, and social status of the individuals to be selected) are chosen to agree with the corresponding proportions in the population. The resulting sample may not be representative of characteristics that have not been taken into account.

QUOTIENT The result of the division of a numerator by a denominator.

R

RACE Persons who are relatively homogeneous with respect to biological inheritance.

See also ETHNIC GROUP.

RADIX The hypothetical size of the birth cohort in a life table, commonly 1000 or 100,000.

RAHE-HOLMES SOCIAL READJUSTMENT RATING SCALE See LIFE EVENTS.

RAMAZZINI, BERNARDINO (1633-1714) An Italian physician, "Father of Occupational Medicine;" he published *De Morbis Artificum* (On the Diseases of Workers) in 1700. Based on observation and anecdote, this was the first systematic account of diseases related to workplace exposures.

RANDOM Governed by chance; not completely determined by other factors. As opposed to deterministic.

RANDOM ALLOCATION See RANDOMIZATION.

RANDOMIZATION Allocation of individuals to groups, e.g., for experimental and control regimens, by chance. Within the limits of chance variation, randomization should make the control and experimental groups similar at the start of an investigation and ensure that personal judgment and prejudices of the investigator do not influence allocation.

Randomization or random assignment should not be confused with haphazard assignment. Random assignment follows a predetermined plan that is usually devised with the aid of a table of random numbers. The pattern of assignment may appear to be haphazard, but this arises from the haphazard nature with which digits occur in a table of random numbers, and not from the haphazard whim of the investigator in allocating patients.

RANDOMIZED CONTROLLED TRIAL (RCT) An epidemiologic experiment in which subjects in a population are randomly allocated into groups, usually called "study" and "control" groups, to receive or not to receive an experimental preventive or therapeutic procedure, maneuver, or intervention. The results are assessed by rigorous comparison of rates of disease, death, recovery, or other appropriate outcome in the study and control groups, respectively. Randomized controlled trials are generally regarded as the most scientifically rigorous method of hypothesis testing available in epidemiology. A few authors refer to this method as "randomized control trial." See also EXPERIMENTAL EPIDEMIOLOGY.

RANDOM SAMPLE A sample that is arrived at by selecting sample units such that each possible unit has a fixed and determinate probability of selection. See also SAMPLE.

RANGE OF DISTRIBUTION The difference between the largest and smallest values in a distribution.

RANKING SCALE (Ordinal Scale) A scale that arrays the members of a group from high to low according to the magnitude of the observations, assigns numbers to the ranks, and neglects distances between members of the array.

RATE A rate is a measure of the frequency of a phenomenon. In epidemiology, demography, and vital statistics, a rate is an expression of the frequency with which an event occurs in a defined population; the use of rates rather than raw numbers is essential for comparison of experience between populations at different times, different places, or among different classes of persons.

The components of a rate are the numerator, the denominator, the specified time in which events occur, and usually a multiplier, a power of 10, which converts the rate from an awkward fraction or decimal to a whole number:

$$\text{Rate} = \frac{\text{Number of events in specified period}}{\text{Average population during the period}} \times 10^n$$

All rates are ratios, calculated by dividing a numerator, e.g., the number of deaths, or newly occurring cases of a disease in a given period, by a denominator, e.g., the average population during that period. Some rates are proportions, i.e., the numerator is contained within the denominator. Rate has several different usages in epidemiology.

1. As a synonym for ratio, it refers to proportions as rates, as in the terms cumulative incidence rate, prevalence rate, survival rate (cf. Webster's Dictionary, which gives proportion and ratio as synonyms for rate).
2. In other situations, rate refers only to ratios representing relative changes (actual or potential) in two quantities. This accords with the OED, which gives "relative amount of variation" among its entries for rate.
3. Sometimes rate is further restricted to refer only to ratios representing changes over time. In this usage, prevalence rate would not be a "true" rate because it cannot be expressed in relation to units of time but only to a "point" in time; in contrast, the force of mortality or force of morbidity (hazard rate) is a "true" rate for it can be expressed as the number of cases developing per unit time, divided by the total size of the population at risk.

RATE DIFFERENCE (RD) The absolute difference between two rates, for example, the difference in incidence rate between a population group exposed to a causal factor and a population group not exposed to the factor:

$$RD = I_e - I_u$$

where I_e = incidence rate among exposed, and I_u = incidence rate among unexposed. In comparisons of exposed and unexposed groups, the term excess rate may be used as a synonym for rate difference.

RATE-ODDS RATIO See ODDS RATIO.

RATE RATIO (RR) The ratio of two rates. The term is used in epidemiologic research with a precise meaning, i.e., the ratio of the rate in the exposed population to the rate in the unexposed population:

$$RR = \frac{I_e}{I_u}$$

where I_e is the incidence rate among exposed, and I_u is the incidence rate among unexposed. See also RELATIVE RISK.

RATIO The value obtained by dividing one quantity by another: a general term of which rate, proportion, percentage, etc., are subsets. The important difference between a proportion and a ratio is that the numerator of a proportion is included in the population defined by the denominator, whereas this is not necessarily so for a

ratio. A ratio is an expression of the relationship between a numerator and a denominator where the two usually are separate and distinct quantities, neither being included in the other.

The dimensionality of a ratio is obtained through algebraic cancellation, summation, etc., of the dimensionalities of its numerator and denominator terms. Both counted and measured values may be included in the numerator and in the denominator. There are no general restrictions on the dimensionalities or ranges of ratios, as there are in some of its subsets (e.g., proportion, prevalence). Ratios are sometimes expressed as percentages (e.g., standardized mortality ratio, FEV₁, percent). In these cases, unlike the special case of a PROPORTION, the value may exceed 100. See also PROPORTION; RATE.

RATIO SCALE See MEASUREMENT SCALE.

RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE (Syn: relative operating characteristic curve) A graphic means for assessing the ability of a screening test to discriminate between healthy and diseased persons. The term "receiver operating characteristic" comes from psychometry where the characteristic operating response of a receiver-individual to faint stimuli or nonstimuli was recorded.

RECORD LINKAGE A method for assembling the information contained in two or more records, e.g., in different sets of medical charts, and in vital records such as birth and death certificates, and a procedure to ensure that the same individual is counted only once. This procedure incorporates a unique identifying system such as a personal identification number and/or birth name(s) of the individual's mother.

Record linkage makes it possible to relate significant health events that are remote from one another in time and place or to bring together records of different individuals, e.g., members of a family. The resulting information is generally stored and retrieved by computer, which can be programmed to tabulate and analyze the data.

"Each person in the world creates a book of life. This book starts with birth and ends with death. Its pages are made of the records of the principal events in life. Record linkage is the name given to the process of assembling the pages of this book into a volume."

¹Dunn HL: Record linkage. *Am J Pub Health* 36:1412, 1946

RECRUDESCENCE Reactivation of infection.

REED, WALTER (1851–1902) US Army physician and epidemiologist. Responsible for epidemiologic investigations and experiments that established the transmission of yellow fever by a filterable virus carried by culicine mosquitoes. The rigorous logic applied to both the experimental and incidental observations by Reed and his colleagues is recognized as one of the great achievements of medical science.

REFERENCE POPULATION The standard against which a population that is being studied can be compared.

REFINEMENT The process of identifying new subcategories of study variables for the purpose of more accurate or more detailed description of relationships. An example is refinement of the concept of serum cholesterol level into high, low, and very low density lipoproteins.

REGISTER, REGISTRY In epidemiology the term "register" is applied to the file of data concerning all cases of a particular disease or other health-relevant condition in a defined population such that the cases can be related to a population base. With this information incidence rates can be calculated. If the cases are regularly followed up, information on remission, exacerbation, prevalence, and survival can also be obtained. The register is the actual document, and the registry is the system of ongoing registration.

In most developed countries all births and deaths are recorded through birth and death registration systems. Results and summaries are then tabulated and published. Examples of registries that have epidemiologic value include the following:

Cancer registries, which secure reports of cancer patients as soon as possible after first diagnosis. The principal sources for these reports are the hospitals serving the community, but a few cases are not reported until death.

Tumor registries, which have provided the basis for studies attempting to differentiate genetic from environmental factors in the etiology of cancer, and other conditions where both genetic and environmental factors may be contributing causes.

Birth defect registries, which seek to document anomalies that are apparent at or soon after birth. They suffer from incompleteness due to omission of stillbirths and of anomalies that do not declare their presence until later in life, such as certain forms of congenital heart lesion, mental deficiency, and neurological disorders.

Other types of registers include blindness and other forms of physical handicap, high-risk infants, persons addicted to drugs, etc. Most of these, however, are not truly population based, but merely list those persons known to or attending some agency or service that provides for them.

REGISTRATION The term "registration" implies something more than notification for the purpose of immediate action or to permit the counting of cases. A register requires that a permanent record be established, including identifying data. Cases may be followed up, and statistical tabulations may be prepared both on frequency and on survival. In addition, the persons listed on a register may be subjects of special studies.

REGRESSION

- As used by FRANCIS GALTON, regression meant the tendency for offspring of exceptional parents (very tall, very intelligent, etc.) to possess characteristics closer to the average for the general population. (Hence, "regression to the mean.")
- In statistics, regression is a synonym for REGRESSION ANALYSIS.

REGRESSION ANALYSIS Given data on a dependent variable *y* and one or more independent variables *x*₁, *x*₂, etc. regression analysis involves finding the "best" mathematical model (within some restricted class of models) to describe *y* as a function of the *x*'s, or to predict *y* from the *x*'s. The most common form is a linear model; in epidemiology, the logistic and proportional hazards models are also common.

REGRESSION LINE Diagrammatic presentation of a regression equation, usually drawn with the independent variable, *x*, as the abscissa and the dependent variable, *y*, as ordinate. Three variables can be shown diagrammatically on an isometric chart or stereogram.

RELATIONSHIP See ASSOCIATION.

RELATIVE ODDS See ODDS RATIO.

RELATIVE RISK

- The ratio of the risk of disease or death among the exposed to the risk among the unexposed; this usage is synonymous with RISK RATIO.
- Alternatively, the ratio of the cumulative incidence rate in the exposed to the cumulative incidence rate in the unexposed, i.e., the cumulative incidence ratio.
- The term "relative risk" has also been used synonymously with "odds ratio" and, in some biostatistical articles, has been used for the ratio of FORCES of MORBIDITY. The use of the term "relative risk" for several different quantities arises from the fact that for "rare" diseases (e.g., most cancers) all the quantities approximate one another. For common occurrences (e.g., neonatal mor-

tality in infants under 1500-g birth weight), the approximations do not hold.

See also CUMULATIVE INCIDENCE RATIO; ODDS RATIO; RATE RATIO; RISK RATIO.

RELIABILITY The degree of stability exhibited when a measurement is repeated under identical conditions. Reliability refers to the degree to which the results obtained by a measurement procedure can be replicated. Lack of reliability may arise from divergences between observers or instruments of measurement or instability of the attribute being measured. See also MEASUREMENT, PROBLEMS WITH TERMINOLOGY; OBSERVER VARIATION.

REPEATABILITY (Syn: reproducibility) A test or measurement is repeatable if the results are identical or closely similar each time it is conducted. See also MEASUREMENT, PROBLEMS WITH TERMINOLOGY; RELIABILITY.

REPLACEMENT LEVEL FERTILITY The level of fertility at which a cohort of women are having only enough daughters to replace themselves in the population. By definition, replacement level fertility is equal to a net reproduction rate of 1.00. The total fertility rate is also used as a measure of replacement level fertility; in the United States today, a total fertility rate of 2.12 is considered to be replacement level; it is higher than 2 because of mortality and because of a sex ratio greater than 1 at birth. The higher the mortality rate, the higher is replacement level fertility.

REPLICATION The execution of an experiment or survey more than once so as to confirm the findings, increase precision, and obtain a closer estimation of sampling error. Exact replication should be distinguished from consistency of results on replication. Exact replication is often possible in the physical sciences, but in the biological and behavioral sciences, to which epidemiology belongs, consistency of results on replication is often the best that can be attained. Consistency of results on replication is perhaps the most important criterion in judgments of causality.

REPRESENTATIVE SAMPLE The term "representative" as it is commonly used is undefined in the statistical or mathematical sense; it means simply that the sample resembles the population in some way.

The use of probability sampling will not ensure that any single sample will be "representative" of the population in all possible respects. If, for example, it is found that the sample age distribution is quite different from that of the population, it is possible to make corrections for the known differences. A common fallacy lies in the unwarranted assumption that, if the sample resembles the population closely on those factors that have been checked, it is "totally representative" and that no difference exists between the sample and the universe or reference population.

Kendall and Buckland¹ comment as follows: "In the widest sense, a sample which is representative of a population. Some confusion arises according to whether 'representative' is regarded as meaning 'selected by some process which gives all samples an equal chance of appearing to represent the population'; or, alternatively, whether it means 'typical in respect of certain characteristics, however chosen'. On the whole, it seems best to confine the word 'representative' to samples which turn out to be so, however chosen, rather than apply it to those chosen with the object of being representative."

¹ Kendall MG, Buckland WR: *A Dictionary of Statistical Terms*, 4th ed. London: Longman, 1982.

REPRODUCIBILITY See REPEATABILITY.

REPRODUCTIVE ISOLATION Absence of interbreeding between populations.

RESEARCH DESIGN The procedures and methods, predetermined by an investigator, to be adhered to in conducting a research project.

RESERVOIR OF INFECTION

1. Any person, animal, arthropod, plant, soil, or substance, or a combination of

these, in which an infectious agent normally lives and multiplies, on which it depends primarily for survival, and where it reproduces itself in such a manner that it can be transmitted to a susceptible host.

2. The natural habitat of the infectious agent.

RESOLUTION (Syn: resolving power) A component of a measuring instrument that helps determine precision. The degree of refinement of the measuring process is commonly referred to as the "resolution" or the "resolving power of the system." See also POWER. The capability of distinguishing between things that are indeed separate or distinct from one another.

RESOLVING POWER The capacity of a system to distinguish between truly distinct things that are close together.

RESPONSE RATE The number of completed or returned survey instruments (questionnaires, interviews, etc.) divided by the total number of persons who would have been surveyed if all had participated. Usually expressed as a percentage. Nonresponse can have several causes, e.g., death, removal out of the survey community, and refusal. See also BIAS; COMPLETION RATE; NONPARTICIPANTS.

RETROACTIVE Pertaining to data gathered from medical records or other sources, when data collection took place without prior planning for the needs of an investigation. See also PROTECTIVE; term in limited use.

RETROSPECTIVE STUDY A research design that is used to test etiologic hypotheses in which inferences about exposure to the putative causal factor(s) are derived from data relating to characteristics of the persons under study or to events or experiences in their past. The essential feature is that some of the persons under study have the disease or other outcome condition of interest, and their characteristics and past experiences are compared with those of other, unaffected persons. Persons who differ in the severity of the disease may also be compared. There is disagreement among epidemiologists as to the desirability of using the term "retrospective study" rather than "case control study" to describe this method. See also CASE CONTROL STUDY.

RETROVIRUS This name is given to a family of RNA viruses characterized by the presence of an enzyme, reverse transcriptase, that enables transcription of RNA to DNA inside an affected cell. Thus, retroviruses can make copies of themselves in host cells. The most important retrovirus is the human immunodeficiency virus (HIV); this makes copies of itself in host cells such as T4 "helper" lymphocytes and normal immune responses are disrupted.

RISK The probability that an event will occur, e.g., that an individual will become ill or die within a stated period of time or age. Also, a nontechnical term encompassing a variety of measures of the probability of a (generally) unfavorable outcome. See also PROBABILITY.

RISK ASSESSMENT The qualitative or quantitative estimation of the likelihood of adverse effects that may result from exposure to specified health hazards or from the absence of beneficial influences.

RISK BENEFIT ANALYSIS The process of analyzing and comparing on a single scale the expected positive (benefits) and negative (risks, costs) results of an action, or lack of an action.

RISK BENEFIT RATIO The results of a risk benefit analysis, expressed as the ratio of risks to benefits.

RISK DIFFERENCE (Syn: excess risk) The absolute difference between two risks.

RISK FACTOR An aspect of personal behavior or lifestyle, an environmental exposure, or an inborn or inherited characteristic, which on the basis of epidemiologic evi-

dence is known to be associated with health-related condition(s) considered important to prevent. The term "risk factor" is rather loosely used, with any of the following meanings:

1. An attribute or exposure that is associated with an increased probability of a specified outcome, such as the occurrence of a disease. Not necessarily a causal factor. A **RISK MARKER**.
2. An attribute or exposure that increases the probability of occurrence of disease or other specified outcome. A **DETERMINANT**.
3. A determinant that can be modified by intervention, thereby reducing the probability of occurrence of disease or other specified outcomes. To avoid confusion, it may be referred to as a modifiable risk factor.

RISK MANAGEMENT The steps taken to alter, i.e., reduce, the levels of risk to which an individual or a population is subject.

RISK MARKER (Syn: risk indicator) An attribute that is associated with an increased probability of occurrence of a disease or other specified outcome and that can be used as an indicator of this increased risk. Not necessarily a causal factor. See also **RISK FACTOR**.

RISK RATIO The ratio of two risks.

ROBUST A statistical test or procedure is said to be robust if it is not very sensitive to departures from the assumptions on which it is strictly predicted (e.g., that the data are normally distributed).

Ross, RONALD (1857–1932) Continued in India the work begun by Laveran and Manson on mosquitoes as vectors of infectious disease. In a series of experiments and microscopic dissections, he concluded that only the anopheles mosquitoes carried the malaria parasite and that a developmental stage of the parasite took place in the mosquito (On some peculiar pigmented cells found in two mosquitoes fed on malarial blood *Brl Med J* 1786–1787, 1897). Awarded the Nobel prize for medicine in 1902.

RUBRIC Section or chapter heading. Used in epidemiology with reference to groups of diseases, e.g., as in the **INTERNATIONAL CLASSIFICATION OF DISEASE** (ICD).

S

SAFETY FACTOR A multiplicative factor incorporated in risk assessments or safety standards to allow for unpredictable types of variation, such as variability from test animals to humans, random variation within an experiment, and person-to-person variability. Safety factors are often in the range of 10 to 1000.

SAFETY STANDARDS Under the requirements of the Occupational Safety and Health Act (OSHA, 1970), "occupational safety and health standard" means a standard that requires conditions, or the adoption of one or more practices, means, methods, operations, or processes reasonably necessary or appropriate to provide safe or healthful employment and places of employment. Safety standards may be adopted by national consensus or established by federal regulation. These standards have been adopted in many other nations besides the United States, although some European and other countries have their own standards, which may be lower or higher than those in the United States.

There are several varieties of safety standards:

1. OSHA-promulgated, mainly for carcinogens, also for cotton dust and lead. These are Permissible Exposure Limits (PELs).
2. National Institute of Occupational Safety and Health (NIOSH) recommendations, often lower limits, based on animal toxicity tests, empirical observations, epidemiologic investigations; these are Recommended Exposure Limits (RELs).
3. An older-established set of criteria has been set by the American Conference of Governmental Industrial Hygienists; these are Threshold Limit Values (TLVs) that have now replaced an earlier set of Maximum Allowable Concentrations (MACs).

SAMPLE A selected subset of a population. A sample may be random or nonrandom and may be representative or nonrepresentative. Several types of sample can be distinguished, including the following:

Cluster sample: Each unit selected is a group of persons (all persons in a city block, a family, etc.) rather than an individual.

Grab sample (Syn: sample of convenience): These ill-defined terms describe samples selected by easily employed but basically nonprobabilistic methods. "Man-in-the-street" surveys and a survey of blood pressure among volunteers who drop in at an examination booth in a public place are in this category. It is improper to generalize from the results of a survey based upon such a sample for there is no way of knowing what sorts of bias may have been operating. See also **BIAS**.

Probability (random) sample: All individuals have a known chance of selection. They may all have an equal chance of being selected, or, if a stratified sampling method is used, the rate at which individuals from several subsets are sampled can be varied so as to produce greater representation of some classes than of others.

A probability sample is created by assigning an identity (label, number) to all individuals in the "universe" population, e.g., by arranging them in alphabetical order and numbering in sequence, or simply assigning a number to each, or by grouping according to area of residence and numbering the groups. The next step

2503012447

is to select individuals (or groups) for study by a procedure such as use of a table of random numbers (or comparable procedure) to ensure that the chance of selection is known.

Simple random sample: In this elementary kind of sample each person has an equal chance of being selected out of the entire population. One way of carrying out this procedure is to assign each person a number, starting with 1, 2, 3, and so on. Then numbers are selected at random, preferably from a table of random numbers, until the desired sample size is attained.

Stratified random sample: This involves dividing the population into distinct subgroups according to some important characteristic, such as age or socioeconomic status, and selecting a random sample out of each subgroup. If the proportion of the sample drawn from each of the subgroups, or strata, is the same as the proportion of the total population contained in each stratum (e.g., age group 40–59 constitutes 20% of the population, and 20% of the sample comes from this age stratum), then all strata will be fairly represented with regard to numbers of persons in the sample.

Systematic sample: The procedure of selecting according to some simple, systematic rule, such as all persons whose names begin with specified alphabetic letters, born on certain dates, or located at specified points on a master list. A systematic sample may lead to errors that invalidate generalizations. For example, persons' names more often begin with certain letters of the alphabet than with other letters, e.g., q, x. A systematic alphabetical sample is therefore likely to be biased.

SAMPLE, EPSEM ("equal probability of selection method") A sample selected in such a manner that all the population units have the same probability of selection. A simple random sample is an Epsem sample; a stratified sample is not unless the probability of selection is the same for all strata.

SAMPLING The process of selecting a number of subjects from all the subjects in a particular group or "universe." Conclusions based on sample results may be attributed only to the population sampled. Any extrapolation to a larger or different population is a judgment or a guess and is not part of statistical inference.

SAMPLING ERROR See **ERROR**.

SAMPLING VARIATION Since the inclusion of individuals in a sample is determined by chance, the results of analysis in two or more samples will differ, purely by chance. This is known as "sampling variation."

SANITARY CORDON See **CORDON SANITAIRE**.

SCATTER DIAGRAM (Syn: scattergram) A graphic method of displaying the distribution of two variables in relation to each other. The values for one variable are measured on the horizontal axis and the values for the other on the vertical axis.

SCENARIO BUILDING A method of predicting the future that relies on a series of assumptions about alternative possibilities, rather than on simple extrapolation of existing trends. Trend lines for demographic composition, morbidity and mortality rates, etc., can then be modified by allowing for each assumption in turn, or combinations of assumptions. The method is claimed to lead to greater flexibility in long-range health planning than simple forecasting that relies only upon extrapolation of trends.

SCREENING Screening was defined in 1951 by the US Commission on Chronic Illness as, "The presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment."

2503012448

Screening is an initial examination only, and positive responders require a second, diagnostic examination. The initiative for screening usually comes from the investigator or the person or agency providing care rather than from a patient with a complaint. Screening is usually concerned with chronic illness and aims to detect disease not yet under medical care.

There are different types of medical screening, each with its own aim: mass, multiple or multiphasic, and prescriptive.

Mass screening simply means the screening of a whole population.

Multiple or multiphasic screening involves the use of a variety of screening tests on the same occasion.

Prescriptive screening has as its aim the early detection in presumptively healthy individuals of disease that can be controlled better if detected early in its natural history.

The characteristics of a screening test include accuracy, estimates of yield, precision, reproducibility, sensitivity and specificity, and validity. See entries under these headings.

SCREENING LEVEL The normal limit or cutoff point at which a screening test is regarded as positive.

SEASONAL VARIATION Change in physiological status or in disease occurrence that conforms to a regular seasonal pattern.

SECONDARY ATTACK RATE The proportion of contacts who get a communicable disease as a consequence of contact with a case. The secondary attack rate is a measure of contagiousness and is useful in evaluating control measures. See also **ATTACK RATE**.

SECULAR TREND (Syn: temporal trend) Changes over a long period of time, generally years or decades. Examples include the decline of tuberculosis mortality and the rise, followed by a decline, in coronary heart disease mortality in the United States and many other countries in the past 50 years.

SELECTION In genetics, the force that brings about changes in the frequency of alleles and genotypes in populations through differential reproduction. In epidemiology, the process and procedure for choosing individuals for study, usually by an orderly means such as random allocation.

SELECTION BIAS See **BIAS**.

SEMMELWEIS, IGNAZ PHILIPP (1818–1865) An Austro-Hungarian physician-obstetrician, who discovered the cause of puerperal fever by carefully comparing infection rates in two wards of the *Allgemeines Krankenhaus* in Vienna. In one ward students customarily came direct from the mortuary or the dissecting room to the patients' bedside whereas in the other, they did not. Puerperal infection death rates were much greater in the former. Semmelweis concluded that some morbid factor was thus transmitted to women in the worse-affected ward. Unhappily, his conclusions were rejected by his colleagues.

SENSITIVITY AND SPECIFICITY (of a screening test) **Sensitivity** is the proportion of truly diseased persons in the screened population who are identified as diseased by the screening test. Sensitivity is a measure of the probability of correctly diagnosing a case, or the probability that any given case will be identified by the test (Syn: true positive rate).

Specificity is the proportion of truly nondiseased persons who are so identified by the screening test. It is a measure of the probability of correctly identifying a non-diseased person with a screening test (Syn: true negative rate). The relationships are shown in the following fourfold table, in which the letters *a*, *b*, *c*, and *d* represent the quantities specified below the table.

| Activity testing | | 120 | | TOTAL |
|------------------------|-------------|----------|--------------|---------|
| Screening test results | True status | Diseased | Not diseased | |
| Positive | | a | b | a+b |
| Negative | | c | d | c+d |
| Total | | a+c | b+d | a+b+c+d |

- a. Diseased individuals detected by the test (true positives)
- b. Nondiseased individuals positive by the test (false positives)
- c. Diseased individuals not detectable by the test (false negatives)
- d. Nondiseased individuals negative by the test (true negatives)

$$\text{Sensitivity} = \frac{a}{a+c} \quad \text{Specificity} = \frac{d}{b+d}$$

$$\text{Predictive value (positive test result)} = \frac{a}{a+b}$$

$$\text{Predictive value (negative test result)} = \frac{d}{c+d}$$

See also YOUDEN'S TEST.

SENSITIVITY TESTING A study of how the final outcome of an analysis changes as a function of varying one or more of the input parameters in a prescribed manner.

SENTINEL HEALTH EVENT A condition that can be used to assess the stability or change in health levels of a population, usually by monitoring mortality statistics. Thus, death due to acute head injury is a sentinel event for a class of severe traffic injury that may be reduced by such preventive measures as use of seatbelts and crash helmets.

SENTINEL PHYSICIAN, SENTINEL PRACTICE In family medicine, a physician, practice, that undertakes to maintain surveillance for and report certain specific predetermined events, such as cases of certain communicable diseases, adverse drug reactions.

SEQUENTIAL ANALYSIS A statistical method that allows an experiment to be ended as soon as an answer of the desired precision is obtained. Study and control subjects are randomly allocated in pairs or blocks. The result of the comparison of each pair of subjects, one treated and one control, is examined as soon as it becomes available and is added to all previous results.

SERENDIPITY The accidental (and happy) discovery of important new information. A well-known example is Fleming's discovery of the bacteriocidal properties of penicillin mould. In case-control studies aimed at testing a specific hypothesis, e.g., about the relationship between tobacco and cancer, questions on other aspects of life-style have serendipitously revealed statistically significant associations, e.g., between alcohol consumption and certain cancers.

SEROEPIDEMIOLOGY Epidemiologic study or activity based on the detection or serological testing of characteristic change in the serum level of specific antibodies. Latent, subclinical infections and carrier states can thus be detected, in addition to clinically overt cases.

SEX RATIO The ratio of one sex to the other. Usually defined as the ratio of males to females (or of the rates observed in males and females).

"SHOE-LEATHER" EPIDEMIOLOGY Gathering information for epidemiologic studies by direct inquiry among the people, e.g., walking from door to door and asking questions of every householder (wearing out shoe leather in the process). JOHN SNOW did this when investigating the sources of water supply to households in the cholera epidemic in London in 1854; the method has been successfully used in many sub-

sequent epidemic investigations. It is especially useful in transmitted diseases.

SIBLINGS Children borne by the same mother.

SIBSHIP All the brothers and sisters borne by the same mother.

SICKNESS See DISEASE.

SIDE EFFECT An effect, other than the intended one, produced

by a diagnostic, or therapeutic procedure or regimen.

SIGNAL-TO-NOISE RATIO A jargon term for the relationship of which is extraneous or irrelevant, or intrudes because no other procedures are sufficiently sensitive.

SIGNIFICANCE See STATISTICAL SIGNIFICANCE.

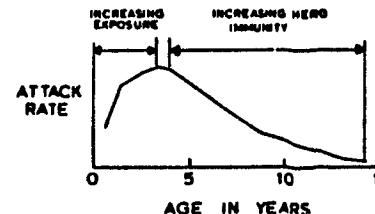
SIMPSON'S PARADOX A form of confounding, in which the presence of a third variable changes the direction of an association. Simpson's paradox is often seen in meta-analysis, because the sum of the data or results from different studies may be affected by confounding variables that have different sign features from some studies but not others; if this analysis will be flawed. Rothman¹ has pointed out that this is really a paradox but the logical consequence of failing to control for confounding variables.

¹Rothman KJ: A pictorial representation of confounding in epidemiology. *Epidemiology* 1975; 6:101-108, 1975.

SIMULATION The use of a model system, e.g., a mathematical model, to approximate the action of a real system, often used to study real system.

SITUATION ANALYSIS Study of a situation that may require intervention. It involves with a definition of the problem, and an assessment of its severity, causes, and impacts upon the community, and is concerned with interactions between the system and its environment and its influence.

SKEW DISTRIBUTION An older and less recommended term for a distribution that is not unimodal. If a unimodal distribution has a long tail extending towards lower values of the variate, it is said to have negative skewness; if it extends towards higher values, it is said to have positive skewness. See also LOG-NORMAL DISTRIBUTION.



Skew distribution of attack rate of measles in relation to age. From Lilienfeld and Lilienfeld, 1979.

SLOW VIRUS Agent causing degenerative (neurological) disease with a long incubation period and a prolonged, slowly progressive course. Some confirmed slow virus diseases are Creutzfeldt-Jakob disease and progressive subcortical gliosis. AIDS is possibly a slow virus disease. Some cases of AIDS are due to slow viruses.

2503012449

SNOW, JOHN (1813–1858) London general practitioner and early anesthetist (he assisted Queen Victoria's delivery of two of her children with chloroform). His fame rests upon his observations, brilliant deductions, painstaking personal enquiries, and analytic studies of cholera outbreaks in the mid-19th century in London and elsewhere. All are recorded in *On the Mode of Communication of Cholera* (London: Churchill, 2nd ed., 1855), which can be regarded as the first definitive working text on epidemiology and which also contained an explicit statement of the germ theory of transmission, written 30 years before Koch discovered the cholera vibrio. See also NATURAL EXPERIMENT.

SOCIAL CLASS A stratum in society composed of individuals and families of equal standing. See also SOCIOECONOMIC CLASSIFICATION.

SOCIAL DRIFT Downward social class mobility as a result of impaired health often due to mental disorders.

SOCIAL MEDICINE The practice of medicine concerned with health and disease as a function of group living. Social medicine is concerned with the health of people in relation to their behavior in social groups and as such involves care of the individual patient as a member of a family and of other significant groups in everyday life. It is also concerned with the health of these groups as such and with that of the whole community as a community. See also COMMUNITY MEDICINE; PUBLIC HEALTH.

SOCIOECONOMIC CLASSIFICATION Arrangement of persons into groups according to such characteristics as prior education, occupation, and income. This usually reveals upon analysis a strong correlation with health-related characteristics such as average length of life and risk of dying from certain specific causes.

The oldest such classification that is epidemiologically useful is the Registrar-General's (RG's) occupational classification, developed in 1911 by Stephenson, Registrar-General of England and Wales. This classified all occupations into five groups—the five "social classes." Social class III is often further subdivided into nonmanual and manual groups:

- I Professional occupations
- II Intermediate occupations
- III Nonmanual skilled occupations
- III M Manual skilled occupations
- IV Partly skilled occupations
- V Unskilled occupations

This has proven to be a valuable epidemiologic tool; social class is an accurate, consistent predictor of health experience.

There have been several other attempts to develop a more refined classification; however, most refinements require collection of more detailed information. For example, Hollingshead's scale requires details about education and income as well as occupation, and so is more time-consuming, more likely to be incomplete, and requires more costly analysis than the RG's classification. In developing countries, where up to 90% of the population may be classified under "agriculturalist" or "pastoralist" (farming or herding), other types of classifications have been developed.

One's prestige in society, and attitudes or values, e.g., setting a high value on getting a good education, are generally an integral part of social class or socioeconomic status. Attitudes toward health are often part of the set of values and may explain part of the observed difference in health between social classes.

SOCIOECONOMIC STATUS (SES) Descriptive term for a person's position in society, which

may be expressed on an ORDINAL SCALE using such criteria as income, educational level attained, occupation, value of dwelling place, etc.

SOFTWARE See COMPUTER.

BOUNDEX CODE A sequence of letters used for recording names phonetically, especially in RECORD LINKAGE.

SOURCE OF INFECTION The person, animal, object, or substance from which an infectious agent passes to a host. Source of infection should be clearly distinguished from source of contamination, such as overflow of a septic tank contaminating a water supply, or an infected cook contaminating a salad. (See RESERVOIR.)¹

¹From *Control of Communicable Disease in Man*, 14th ed. Washington DC: American Public Health Association, 1985.

SPEARMAN'S RANK CORRELATION See CORRELATION COEFFICIENT.

SPECIFICATION

1. The process of selecting a particular functional form or model for the relationships to be analyzed in a study.
2. The process of selecting variables for inclusion in the analysis of an effect or association. This process leads to the identification of MODERATOR VARIABLES and CONFOUNDING VARIABLES. See also STRATIFICATION.

SPECIFICITY (OR A TEST) See SENSITIVITY AND SPECIFICITY.

SPECTRUM OF DISEASE The full range of manifestations of a disease: a vague term, that can mean everything from mild or subclinical or precursor states to fulminating, florid disease, or alternatively the natural history of a disease from onset to resolution.

SPELL OF SICKNESS An episode of sickness with a well-defined onset and termination. As used in the monitoring or surveillance of disease, the spell is often defined by the duration of absence from work or school.

SPLEEN RATE A term used in malaria epidemiology, to define the frequency of enlarged spleens detected on survey of a population in which malaria is prevalent. In association with the HACKETT SPLEEN CLASSIFICATION it summarizes the severity of malaria endemicity.

SPORADIC Occurring irregularly, haphazardly from time to time, and generally infrequently, e.g., cases of certain infectious diseases.

SPOT MAP Map showing the geographic location of people with a specific attribute, e.g., cases of a disease or elderly persons living alone. The making of a spot map is a common procedure in the investigation of a localized outbreak of disease. Inferences from such a map depend on the assumption that the population at risk of developing the disease is fairly evenly distributed over the area, or that at least the heterogeneities are known and can be considered in interpreting the map.

STABLE POPULATION A population that has constant fertility and mortality rates, no migration, and consequently a fixed age distribution and constant growth rate. See also STATIONARY POPULATION.

STANDARD Something that serves as a basis for comparison; a technical specification or written report drawn up by experts based on the consolidated results of scientific study, technology, and experience, aimed at optimum benefits and approved by recognized and representative body.

STANDARD DEVIATION A measure of dispersion or variation. It is the most widely used measure of dispersion of a frequency distribution. It is equal to the positive square root of the variance. The mean tells where the values for a group are centered. The standard deviation is a summary of how widely dispersed the values are around this center.

STANDARD ERROR The standard deviation of an estimate.

STANDARDIZATION A set of techniques used to remove as far as possible the effects of differences in age or other confounding variables, when comparing two or more populations. The common method uses weighted averaging of rates specific for age, sex, or some other potential confounding variable(s) according to some specified distribution of these variables. There are two main methods, as follows:

Direct method: The specific rates in a study population are averaged, using as weights the distribution of a specified standard population. The directly standardized rate represents what the crude rate would have been in the study population if that population had the same distribution as the standard population with respect to the variable(s) for which the adjustment or standardization was carried out.

Indirect method: This is used to compare study populations for which the specific rates are either statistically unstable or unknown. The specific rates in the standard population are averaged, using as weights the distribution of the study population. The ratio of the crude rate for the study population to the weighted average so obtained is the standardized mortality (or morbidity) ratio, or SMR. The indirectly standardized rate itself is the product of the SMR and the crude rate for the standard population.

STANDARDIZED MORTALITY (MORBIDITY) RATIO (SMR) The ratio of the number of events observed in the study group or population to the number that would be expected if the study population had the same specific rates as the standard population, multiplied by 100.

STANDARDIZED RATE RATIO (SRR) A rate ratio in which the numerator and denominator rates have been standardized to the same (standard) population distribution.

STANDARD METROPOLITAN STATISTICAL AREA Because of the extensive interactions between a city and its surrounding areas, a unit encompassing both is needed as a base for statistical description. The concept of a standard metropolitan statistical area (SMSA) was introduced in the United States to furnish such a unit. To qualify as an SMSA an area has to meet criteria related to size, social and economic integration of the city and surrounding county or counties, minimum population density, and minimum proportion of the labor force engaged in nonagricultural work.

STATIONARY POPULATION A stable population that has a zero growth rate with constant numbers of births and deaths each year.

STATISTICS The science and art of collecting, summarizing, and analyzing data that are subject to random variation. The term is also applied to the data themselves and to summarizations of the data. Statistical terms are defined by Kendall and Buckland.¹

¹Kendall MG, Buckland WR: *A Dictionary of Statistical Terms*, 4th ed. London: Longman, 1982.

STATISTICAL ERROR See **ERROR**.

STATISTICAL INFERENCE See **INFERENCE**.

STATISTICAL MODEL See **MATHEMATICAL MODEL**.

STATISTICAL SIGNIFICANCE Statistical methods allow an estimate to be made of the probability of the observed or greater degree of association between independent and dependent variables under the null hypothesis. From this estimate, in a sample of given size, the statistical "significance" of a result can be stated. Usually the level of statistical significance is stated by the P value.

STATISTICAL TEST A procedure that is intended to decide whether a hypothesis about the distribution of one or more populations or variables should be rejected or accepted. Statistical tests may be parametric or nonparametric.

STEREOGRAM (Syn: isometric chart) A graph or chart that displays more than two variables in a manner that appears three-dimensional to the eye.

STOCHASTIC PROCESS A process that incorporates some element of randomness.

2503012451

STRATEGY In game theory, a mathematical function.

STRATIFICATION The process of or result of separating a sample into several subsamples according to specified criteria such as age groups, socioeconomic status, etc. The effect of confounding variables may be controlled by stratifying the analysis of results. For example, lung cancer is known to be associated with smoking. To examine the possible association between urban atmospheric pollution and lung cancer, controlling for smoking, the population may be divided into strata according to smoking status. The association between air pollution and cancer can then be appraised separately within each stratum. Stratification is used not only to control for confounding effects but also as a way of detecting modifying effects. In this example, stratification makes it possible to examine the effect of smoking on the association between atmospheric pollution and lung cancer.

STRATIFIED RANDOMIZATION (Syn: blocked randomization) A randomization procedure in which strata are identified and subjects randomly allocated within each. This produces a situation intermediate between paired allocation and simple random allocation.

STUDY DESIGN See **RESEARCH DESIGN**.

SUBCLINICAL DISEASE See **DISEASE, SUBCLINICAL**.

SURVEILLANCE Ongoing scrutiny, generally using methods distinguished by their practicability, uniformity, and frequently their rapidity, rather than by complete accuracy. Its main purpose is to detect changes in trend or distribution in order to initiate investigative or control measures. See also **MONITORING**.

SURVEILLANCE OF DISEASE The continuing scrutiny of all aspects of occurrence and spread of a disease that are pertinent to effective control.

Included are the systematic collection and evaluation of (1) morbidity and mortality reports, (2) special reports of field investigations of epidemics and of individual cases, (3) isolation and identification of infectious agents by laboratories, (4) data concerning the availability, use, and untoward effects of vaccines and toxoids, immune globulins, insecticides, and other substances used in control, (5) information regarding immunity levels in segments of the population, and (6) other relevant epidemiologic data. A report summarizing these data should be prepared and distributed to all cooperating persons and others with a need to know the results of the surveillance activities. The procedure applies to all jurisdictional levels of public health from local to international.¹ Serological surveillance identifies patterns of current and past infection using serological test. See also **SEROEPIDEMIOLOGY**.

¹Benenson AS (Ed.): *Control of Communicable Diseases in Man*, 14th ed. Washington DC: American Public Health Association, 1985.

SURVEY An investigation in which information is systematically collected but in which the experimental method is not used. A population survey may be conducted by face-to-face inquiry, by self-completed questionnaires, by telephone, postal service, or in some other way. Each method has its advantages and disadvantages. For instance, a face-to-face (interview) survey may be a better way than self-completed questionnaire to collect information on attitudes or feelings, but it is more costly. Existing medical or other records may contain accurate information, but not about a representative sample of the population.

The information that is gathered in a survey is usually complex enough to require editing (for accuracy, completeness, etc.), coding, keypunching, i.e., entry on **PUNCH CARDS** and processing and analysis by machine or computer. The generalizability of results depends upon the extent to which the surveyed population is representative.

The term "survey" is sometimes used in a narrow sense to refer specifically to a FIELD SURVEY.

SURVEY INSTRUMENT The interview schedule, questionnaire, medical examination record form, etc., used in a survey.

SURVIVAL ANALYSIS A class of statistical procedures for estimating the SURVIVAL FUNCTION, and for making inferences about the effects on it of treatments, prognostic factors, exposures, and other covariates.

SURVIVAL CURVE A curve that starts at 100% of the study population and shows the percentage of the population still surviving at successive times for as long as information is available. May be applied not only to survival as such, but also to the persistence of freedom from a disease, or complication or some other endpoint.

SURVIVAL FUNCTION (Svn: survival distribution) A function of time, usually denoted by $S(t)$, that starts with a population 100% well at a particular time and provides the percentage of the population still well at later times. Survival functions may be applied to any discrete event, for example, disease incidence or relapse, death, or recovery after onset of disease (in which case the population is initially 100% diseased, and the "survival" function gives the percentage still diseased).

SURVIVAL RATE (Svn: cumulative survival rate) The proportion of survivors in a group, e.g., of patients, studied and followed over a period. The proportion of persons in a specified group alive at the beginning of the time interval (e.g., a five-year period) who survive to the end of the interval. It is equal to 1 minus the CUMULATIVE MORTALITY RATE. May be studied by current or COHORT LIFE TABLE methods.

SURVIVAL RATIO The probability of surviving between one age and another; when computed for age groups, the ratios correspond to those of the person-years-lived function of a life table.

SURVIVORSHIP STUDY Use of a cohort LIFE TABLE to provide the probability that an event, such as death, will occur in successive intervals of time after diagnosis and, conversely, the probability of surviving each interval. The multiplication of these probabilities of survival for each time interval for those alive at the beginning of that interval yields a cumulative probability of surviving for the total period of study.

SYDENHAM, THOMAS (1624–1689) A great English physician in the tradition of Hippocrates and one of the founding fathers of epidemiology (although his ideas about the meteorological causes of epidemics were wrong). His writings contain many careful and comprehensive accounts of important epidemic diseases, notably plague, malaria, measles, dysentery, and scarlet fever. His *Opera Omnia* have been twice translated into English: the second (and better) two-volume translation by Latham was published by the Sydenham Society in 1848–1850.

SYMBIOSIS The biological association of two or more species to their mutual benefit.

SYMMETRICAL RELATIONSHIP An association between variables that does not have direction.

The following four varieties can be distinguished:

1. Functional interdependence, where one variable cannot exist without the other; e.g., prevalence is a function of incidence and duration.
2. Common complex, where variables occur together without being interdependent or necessary to each other; e.g., the occurrence together of air pollution, poverty, poor housing, and overcrowding.
3. Alternative indicators of the same entity; e.g., antibodies to a microorganism and history of specific infection caused by that microorganism.
4. The effects of a common cause; e.g., clinical and biochemical changes in hepatitis.

See also ASSOCIATION, SYMMETRICAL.

2503012452

SYNDROME A symptom complex in which the symptoms and/or signs coexist more frequently than would be expected by chance on the assumption of independence.

SYNERGISM, SYNERGY The definition of synergism in epidemiology is somewhat controversial. We offer two definitions, the first a common dictionary definition, the second a more specific definition encountered in bioassay.

1. A situation in which the combined effect of two or more factors is greater than the sum of their solitary effects.
2. Two factors act synergistically if there are persons who will get the disease when exposed to both factors but not when exposed to either alone. ANTAGONISM, the opposite of synergism, exists if there are persons who will get the disease when exposed to one of the factors alone, but not when exposed to both. Note that under these definitions two factors may act synergistically in some persons and antagonistically in others.

SYSTEMATIC ERROR See BIAS.

SYSTEMS ANALYSIS This term is used with three similar meanings:

1. The examination of various elements of a system with a view to ascertaining whether the proposed solution to a problem will fit into the system and, in turn, effect an overall improvement in the system.
2. The analysis of an activity in order to determine precisely what is required of the system, how this can best be accomplished, and in what ways the computer can be useful.
3. Systems analysis refers to any formal analysis whose purpose is to suggest a course of action by systematically examining the objectives, costs, effectiveness and risks of alternative policies or strategies and designing additional ones if those examined are found wanting. It is an approach to or way of looking at complex problems of choice under uncertainty; it is not yet a method.

T

TAKAKI, KANEHIRO (1849–1915) Japanese nobleman who studied medicine at St Thomas's Hospital Medical School, London. He became a naval surgeon, and later used his opportunity as director of naval medical services to conduct large-scale dietary experiments on populations of naval personnel, demonstrating that beriberi could be prevented by a mixed diet containing protein as well as rice.

TARGET POPULATION

1. The collection of individuals, items, measurements, etc., about which we want to make inferences. The term is sometimes used to indicate the population from which a sample is drawn and sometimes to denote any "reference" population about which inferences are required.
2. The group of persons for whom an intervention is planned.

TAXONOMY A systematic classification into related groups.

TAXONOMY OF DISEASE The orderly classification of diseases into appropriate categories on the basis of relationships among them, with the application of names. See also NOSOGRAPHY, NOSOLOGY.

t-DISTRIBUTION, t-TEST The *t*-distribution is the distribution of a quotient of independent random variables, the numerator of which is a standardized normal variate and the denominator of which is the positive square root of the quotient of a chi-square distributed variate and its number of degrees of freedom. The *t*-test uses a statistic that, under the null hypothesis, has the *t*-distribution, to test whether two means differ significantly, or to test linear regression or correlation coefficients. The *t*-distribution and the *t*-test were developed by WS Gossett, who wrote under the pseudonym "Student" as his employment precluded individual publication.

TERATOGEN A substance that produces abnormalities in the embryo or fetus by disturbing maternal homeostasis or by acting directly on the fetus in utero.

TEST OF SIGNIFICANCE See P VALUE; STATISTICAL SIGNIFICANCE.

TEST HYPOTHESIS See NULL HYPOTHESIS.

THEORETICAL EPIDEMIOLOGY The development of mathematical/statistical models to explain different aspects of the occurrence of a variety of diseases. With some infectious diseases, models have been generated to elucidate the reasons for epidemics and/or to predict the behavior of the disease in reaction to given control measures. See also MODEL.

THERAPEUTIC TRIAL See CLINICAL TRIAL.

THRESHOLD LIMIT VALUE See SAFETY STANDARDS.

THRESHOLD PHENOMENA Events or changes that occur only after a certain level of a characteristic is reached.

TIME CLUSTER See CLUSTERING.

TIME-PLACE CLUSTER See CLUSTERING.

TOTAL FERTILITY RATE (TFR) The average number of children that would be born per woman if all women lived to the end of their childbearing years and bore children according to a given set of age-specific fertility rates. It is computed by summing the age-specific fertility rates for all ages and multiplying by the interval into which the ages are grouped. The TFR is an important fertility measure, providing the most accurate answer to the question, "How many children does a women have, on average?"

TRACER DISEASE METHOD Tracer or indicator conditions as defined by Kessner¹ are easily diagnosed, reasonably frequent illnesses or health states whose outcomes are believed to be affected by health care and which taken in aggregate should reflect the gamut of patients and health problems encountered in a medical practice. The extent to which the recorded care of these conditions concurs with preset standards of care is used as an index of the quality of care delivered. However, it should first be shown that the preset standards contribute to a favorable outcome. See also SENTINEL HEALTH EVENT.

¹Kessner DM, Snow CK, Singer J: *Assessment of Medical Care for Children*. Washington DC: National Academy of Sciences, Institute of Medicine, 1974.

TRANSMISSION OF INFECTION Transmission of infectious agents. Any mechanism by which an infectious agent is spread through the environment or to another person. These mechanisms are defined in *Control of Communicable Disease in Man*¹ as follows:

a. Direct transmission

Direct and essentially immediate transfer of infectious agents (other than from an arthropod in which the organism has undergone essential multiplication or development) to a receptive portal of entry through which human infection may take place. This may be by direct contact as by touching, kissing, or sexual intercourse, or by the direct projection (droplet spread) of droplet spray onto the conjunctiva or onto the mucous membranes of the nose or mouth during sneezing, coughing, spitting, singing, or talking (usually limited to a distance of about 1 m or less). It may also be by direct exposure of susceptible tissue to an agent in soil, compost, or decaying vegetable matter in which it normally leads a saprophytic existence, (e.g., the systemic mycoses), or by the bite of a rabid animal. Transplacental transmission is another form of direct transmission.

b. Indirect transmission

Vehicle-borne—Contaminated materials or objects (fomites) such as toys, handkerchiefs, soiled clothes, bedding, cooking or eating utensils, and surgical instruments or dressings (indirect contact); water, food, milk, biological products including blood, serum, plasma, tissues, or organs; or any substance serving as an intermediate means by which an infectious agent is transported and introduced into a susceptible host through a suitable portal of entry. The agent may or may not have multiplied or developed in or on the vehicle before being transmitted.

Vector-borne—(1) **Mechanical**: Includes simple mechanical carriage by a crawling or flying insect through soiling of its feet or proboscis, or by passage of organisms through its gastrointestinal tract. This does not require multiplication or development of the organism. (2) **Biological**: Propagation (multiplication), cyclic development, or a combination of these (cyclopropagative) is required before the arthropod can transmit the infective form of the agent to man. An incubation period (extrinsic) is required following infection before

the arthropod becomes infective. The infectious agent may be passed vertically to succeeding generations (transovarian transmission); transstadial transmission is its passage from the one stage of the life cycle to another, as nymph to adult. Transmission may be by saliva during biting or by regurgitation or deposition on the skin of feces or other material capable of penetrating subsequently through the bite wound or through an area of trauma from scratching or rubbing. This is transmission by an infected nonvertebrate host and must be differentiated for epidemiologic purposes from simple mechanical carriage by a vector in the role of a vehicle. An arthropod in either role is termed a "vector."

Airborne—The dissemination of microbial aerosols to a suitable portal of entry, usually the respiratory tract. Microbial aerosols are suspensions in the air of particles consisting partially or wholly of microorganisms. Particles in the $1\text{--}5\mu$ range are easily drawn into the alveoli of the lungs and may be retained there; many are exhaled from the alveoli without deposition. They may remain suspended in the air for long periods of time, some retaining and others losing infectivity or virulence. Not considered as airborne are droplets and other large particles that promptly settle out (see Direct transmission, above).

The following are airborne and their mode of transmission is direct:

Droplet nuclei: Usually the small residues that result from evaporation of fluid from droplets emitted by an infected host (see above). Droplet nuclei also may be created purposely by a variety of atomizing devices, or accidentally as in microbiology laboratories or in abattoirs, rendering plants, or autopsy rooms. They usually remain suspended in the air for long periods of time.

Dust: The small particles of widely varying size that may arise from soil (as, for example, fungus spores separated from dry soil by wind or mechanical agitation), clothes, bedding, or contaminated floors.¹ See also ACQUAINTANCE NETWORK; AIR-BORNE INFECTION; CARRIER; COMMON VEHICLE SPREAD; CONTACT; CONTAMINATION; DROPLET NUCLEI.

¹Benenson AS (Ed.): *Control of Communicable Diseases in Man*, 14th ed. Washington DC: American Public Health Association, 1985.

TRANSOVARIAL TRANSMISSION See VECTOR-BORNE INFECTION.

TRANSPORT HOST See PARATENIC HOST.

TREND A long-term movement in an ordered series, e.g., a time series. An essential feature is that the movement, while possibly irregular in the short term, shows movement consistently in the same direction over a long term. The term is also used loosely to refer to an association which is consistent in several samples or strata but is not statistically significant.

TREND LINE That line that best fits the distribution of a set of values plotted on two axes.

TRIAL See CLINICAL TRIAL.

TROHOC STUDY A retrospective case-control study. The term, proposed by AR Feinstein,¹ is the inversion of "cohort;" its use is deprecated by the great majority of epidemiologists.

¹Clin Pharmacol Ther 30:564-577, 1981.

TYPE I ERROR See ERROR.

TYPE II ERROR See ERROR.

TWIN STUDY Method of detecting genetic etiology in human disease. The basic premise of twin studies is that monozygotic twins, being formed by the division of a single

fertilized ovum, carry identical genes, while dizygotic twins, being formed by the fertilization of two ova by two different spermatozoa, are genetically no more similar than two siblings born after separate pregnancies.

TWO-TAIL TEST A statistical significance test based on the assumption that the data are distributed in both directions from some central value(s).

U,V

UNBIASED ESTIMATOR An estimator that for all sample sizes has an expected value equal to the parameter being estimated. If an estimator tends to be unbiased as sample size increases, it is referred to as asymptotically unbiased.

UNDERLYING CAUSE OF DEATH See DEATH CERTIFICATE.

UNDERREPORTING Failure to identify and/or count all cases, leading to reduction of numerator in a rate. See also ERROR.

UTILITY In economics, this means satisfaction derived from obtaining some quantity of a specified article of commerce. When used in decision theory or CLINICAL DECISION ANALYSIS, the meaning is essentially the same, and can be expressed as the usefulness or desirability of an outcome resulting from a decision.

VACCINATION Strictly speaking, vaccination refers to inoculation (from Latin *in oculus*, into a bud) with vaccinia virus against smallpox. Nowadays the word is broadly used synonymously with procedures for immunization against all infectious disease.

VACCINE Immunobiological substance used for active immunization by introducing into the body a live modified, attenuated, or killed inactivated infectious organism or its toxin. The vaccine is capable of stimulating immune response by the host, who is thus rendered resistant to infection. The word "vaccine" was originally applied to the serum from a cow infected with vaccinia virus (cowpox; from Latin *vaccia*, cow); it is now used of all immunizing agents.

VALIDATION The process of establishing that a method is sound.

VALIDITY This term, derived from the Latin *validus*, strong, has several meanings, usually accompanied by a qualifying word or phrase.

VALIDITY, MEASUREMENT An expression of the degree to which a measurement measures what it purports to measure.

Several varieties are distinguished, including construct validity, content validity, and criterion validity (concurrent and predictive validity).

Construct validity: The extent to which the measurement corresponds to theoretical concepts (constructs) concerning the phenomenon under study. For example, if on theoretical grounds, the phenomenon should change with age, a measurement with construct validity would reflect such a change.

Content validity: The extent to which the measurement incorporates the domain of the phenomenon under study. For example, a measurement of functional health status should embrace activities of daily living, occupational, family, and social functioning, etc.

Criterion validity: The extent to which the measurement correlates with an external criterion of the phenomenon under study. Two aspects of criterion validity can be distinguished:

1. **Concurrent validity:** The measurement and the criterion refer to the same point in time. An example would be a visual inspection of a wound for evidence of

infection validated against bacteriological examination of a specimen taken at the same time.

2. **Predictive validity:** The measurement's validity is expressed in terms of its ability to predict the criterion. An example would be an academic aptitude test that was validated against subsequent academic performance.

VALIDITY, STUDY The degree to which the inference drawn from a study, especially generalizations extending beyond the study sample, are warranted when account is taken of the study methods, the representativeness of the study sample, and the nature of the population from which it is drawn. Two varieties of study validity are distinguished:

1. **Internal validity:** The index and comparison groups are selected and compared in such a manner that the observed differences between them on the dependent variables under study may, apart from sampling error, be attributed only to the hypothesized effect under investigation.

2. **External validity (generalizability):** A study is externally valid or generalizable if it can produce unbiased inferences regarding a target population (beyond the subjects in the study). This aspect of validity is only meaningful with regard to a specified external target population. For example, the results of a study conducted using only white male subjects might or might not be generalizable to all human males (the target population consisting of all human males). It is not generalizable to females (the target population consisting of all people). The evaluation of generalizability usually involves much more subject-matter judgment than internal validity.

These epidemiologic definitions of the terms "internal validity" and "external validity" do not correspond exactly to some definitions found in the sociological literature.

VARIABLE Any quantity that varies. Any attribute, phenomenon, or event that can have different values.

VARIABLE, ANTECEDENT A variable that causally precedes the association or outcome under study. See also EXPLANATORY VARIABLE; INDEPENDENT VARIABLE.

VARIABLE, CONFOUNDING See CONFOUNDING.

VARIABLE, CONTROL Independent variable other than the "hypothetical causal variable" that has a potential effect on the dependent variable and is subject to control by analysis.

VARIABLE, DEPENDENT See DEPENDENT VARIABLE.

VARIABLE, DISTORTER A confounding variable that diminishes, masks, or reverses the association under study.

VARIABLE, EXPERIENTIAL See INDEPENDENT VARIABLE.

VARIABLE INDEPENDENT See INDEPENDENT VARIABLE.

VARIABLE, INTERVENING See INTERVENING VARIABLE.

VARIABLE, MANIFESTATIONAL See DEPENDENT VARIABLE.

VARIABLE, MODERATOR See EFFECT MODIFIER.

VARIABLE, PASSENGER See PASSENGER VARIABLE.

VARIABLE, UNCONTROLLED A (potentially) confounding variable that has not been brought under control by design or analysis. See also CONFOUNDING.

VARIANCE A measure of the variation shown by a set of observations, defined by the sum of the squares of deviations from the mean, divided by the number of DEGREES OF FREEDOM in the set of observations.

VARIATE (Syn: random variable) A variable that may assume any of a set of values, each with a preassigned probability (known as its distribution).

VECTOR

- In infectious disease epidemiology, an insect or any living carrier that transports an infectious agent from an infected individual or its wastes to a susceptible individual or its food or immediate surroundings. The organism may or may not pass through a developmental cycle within the vector.
- In statistics, an ordered set of numbers representing the values of a set of variables.

VECTOR-BORNE INFECTION Several classes of vector-borne infections are recognized, each with epidemiologic features that are determined by the interaction between the infectious agent and the human host, on the one hand, and the vector on the other. Therefore, environmental factors such as climatic and seasonal variations influence the epidemiologic pattern by virtue of their effects on the vector and its habits.

The terms used to describe specific features of vector-borne infections are:

Biological transmission: Transmission of the infectious agent to susceptible host by bite of blood-feeding (arthropod) vector as in malaria, or by other inoculation, as in *Schistosoma* infection.

Extrinsic incubation period: Time necessary after acquisition of infection by the (arthropod) vector for the infectious agent to multiply or develop sufficiently so that it can be transmitted by the vector to a vertebrate host.

Hibernation: A possible mechanism by which the infected vector survives adverse cold weather by becoming dormant.

Inapparent infection: Response to infection without developing overt signs of illness. If this is accompanied by viremia or bacteremia in a high proportion of infected animals or persons, the receptor species is well suited as an epidemiologically important host in the transmission cycle.

Mechanical transmission: Transport of the infectious agent between hosts by arthropod vectors with contaminated mouthparts, antennae, or limbs. There is no multiplication of the infectious agent in the vector.

Overwintering: Persistence of the infectious microorganism in the vector for extended periods, such as the cooler winter months, during which the vector has no opportunity to be reinfected or to infect a vertebrate host. Overwintering is an important concept in the epidemiology of vector-borne diseases since the annual recrudescence of viral activity after periods (winter, dry season) adverse to continual transmission depends upon a mechanism for local survival of an infectious microorganism or its reintroduction from outside the endemic area. To some extent, the risk of a summertime epidemic may be determined by the relative success of microorganism survival in the local winter reservoir. Since overwinter survival may in turn depend upon the level of activity of the microorganism during the preceding summer-fall, outbreaks sometimes occur for two or more successive years.

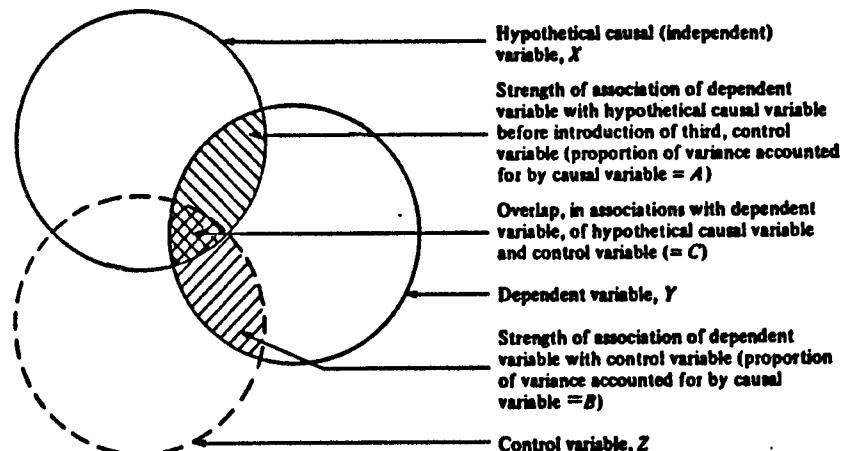
Transovarial infection (transmission): Transmission of the infectious microorganism from the affected female arthropod to her progeny.

VECTOR SPACE An area (or volume) defined by the specified dimensions of two (or three) vectors.

VEHICLE OF INFECTION TRANSMISSION The mode of transmission of an infectious agent from its reservoir to a susceptible host. This can be person-to-person, food, vector-borne, etc.

VENN DIAGRAM A pictorial presentation of the extent to which two or more quantities or concepts are mutually inclusive and mutually exclusive.

VIRCHOW, RUDOLF (1821-1902) Born in Pomerania, Virchow graduated in medicine from Berlin in 1843 and rapidly established his reputation as the leading medical



Venn diagram. From Susser, 1973.

scientist of his time. Modern pathology owes much to his rigorous use of hypothesis-testing methods, illustrated in his first paper in the journal he founded, *Archiv für pathologische Anatomie*, now universally known as Virchow's Archives. Virchow was also a practicing epidemiologist, who investigated a serious epidemic of typhus in Silesia in 1848; his recommendations for hygienic and social reform got him into trouble with the government, but his scientific brilliance made it impossible for the authorities not to recognize and reward him with promotions and honors. He entered Parliament in 1862, and during the Franco-Prussian War he organized an ambulance service. He made many contributions of fundamental importance to the science of pathology, but deserves to be remembered as a great humanitarian as well.

VIRGIN POPULATION A population that has never been exposed to a particular infectious agent.

VIRULENCE The degree of pathogenicity; the disease-evoking power of a microorganism in a given host. Numerically expressed as the ratio of the number of cases of overt infection in the total number infected, as determined by immunoassay. When death is the only criterion of severity, this is the case-fatality rate.

VITAL RECORDS (Literally, "To do with living") Certificates of birth, death, marriage, and divorce required for legal and demographic purposes.

VITAL STATISTICS Systematically tabulated information concerning births, marriages, divorces, separations, and deaths based on registrations of these vital events.

W, X, Y, Z

WASHOUT PHASE That stage in a study, especially a therapeutic trial, when treatment is withdrawn so that its effects disappear and the subject's characteristics return to their baseline state.

WORM COUNT A method of surveillance of helminth infection of the gut that depends upon counts of worms, or their cysts or ova, in quantitatively titrated samples of feces. Other terms used to describe this form of surveillance are "egg count," "cyst count," and "parasite count."

WU, LIEN-TEH (1879–1960) Chinese epidemiologist, responsible for controlling the plague pandemic in Manchuria in 1910–11. Later he worked on control of sexually transmitted diseases and other socioeconomically determined health conditions, developed a national quarantine service and was one of the founders of the Chinese Medical Association, thus helping to lay the foundations for health improvements in modern China.

XENOBIOTIC

1. (Syn: commensal, symbiosis) Pertaining to association of two animal species, usually insects, in the absence of a dependency relationship, as opposed to parasitism.
2. A foreign compound that is metabolized in the body. Many pesticides and their derivatives, some food additives and a number of other complex organic compounds such as dioxins and PCBs, are xenobiotics.

XENODIAGNOSIS Detection of a (human) pathogenic organism by allowing a noninfected vector (e.g., mosquito) to consume infected material, and then examining this vector for evidence of the pathogen.

YATES' CORRECTION An adjustment proposed by Yates (1934) in the chi-square calculation for a 2×2 table, which brings the distribution based on discontinuous frequencies closer to the continuous chi-square distribution from which the published tables for testing chi-squares are derived.

YEARS OF POTENTIAL LIFE LOST (YPLL) See POTENTIAL YEARS OF LIFE LOST.

YIELD The number or proportion of cases of a condition accurately identified by a screening test.

YOUDEN'S INDEX When assessing screening tests, in the uncommon case where the risk of a false negative and that of a false positive result are assumed to be equivalent (i.e., specificity and sensitivity assumed to be equally important), it may be possible to compare screening tests through the Youden index based on the sum of specificity and sensitivity:

$$\text{Youden Index} = J = \text{specificity} + \text{sensitivity} - 1$$

with J ranging from zero (specificity = 0.50 and sensitivity = 0.50) to 1 (sensitivity = 1.00, specificity = 1.00).

ZERO-TIME SHIFT This concerns the selection of a starting point for the measurement of survival following the detection of disease. It is a jargon term, denoting the movement "backward" (toward the starting point of a disease) of time between onset and detection, that may accompany use of a screening procedure.

ZOONOSIS An infection or infectious disease transmissible under natural conditions from vertebrate animals to man. Examples include rabies and plague. May be enzootic or epizootic.

Bibliography

Many of the works on this list contain glossaries, and nearly all contain definitions that have been adapted and included in this dictionary.

Abramson JH: *Survey Methods in Community Medicine*, 3rd ed. London: Churchill Livingstone, 1984.

Alderson M: *An Introduction to Epidemiology*, 2nd ed. London: Macmillan, 1983.

Allaby M: *Dictionary of the Environment*. Southampton: London Press, 1975.

Armitage P: *Statistical Methods in Medical Research*. Oxford: Blackwell, 1971.

Bahn AK: *Basic Medical Statistics*. New York: Grune & Stratton, 1972.

Barker DJP, Rose G: *Epidemiology in Medical Practice*, 2nd ed. Edinburgh: Churchill Livingstone, 1979.

Benenson AS (Ed): *Control of Communicable Diseases in Man*, 14th ed. Washington DC: American Public Health Association, 1985.

Bogue DJ: *Principles of Demography*. New York: Wiley, 1969.

Breslow NE, Day NE: *Statistical Methods in Cancer Research, Vol 1: The Analysis of Case-Control Data*. Lyon: IARC, 1980.

Cavalli-Sforza LL, Bodmer WF: *The Genetics of Human Populations*. San Francisco: Freeman, 1971.

Colton T: *Statistics in Medicine*. Boston: Little, Brown, 1974.

Committee on Population and Demography, National Academy of Science/National Research Council: *Collecting Data for the Estimation of Fertility and Mortality*. Washington DC: National Academy Press, Report Number 6, 1981.

Cox DR, Hinkley DV: *Theoretical Statistics*. New York: Chapman and Hall, 1974.

Davies G: *A Dictionary of Veterinary Epidemiology*. Mimeographed. To be published.

Dublin LI, Lotka AJ: *The Money Value of a Man*. New York: Ronald, 1930.

Farr, W: *Vital Statistics* (Ed. NA Humphreys). London: Stanford, 1885. [A modern abridgement, edited by A Adelstein and MW Susser, was published by the New York Academy of Medicine in 1975.]

Feinstein AR: *Clinical Biostatistics*. St Louis: Mosby, 1974.

Feinstein AR: *Clinical Epidemiology*. Philadelphia: Saunders, 1985.

Feinstein AR: A glossary of neologisms in quantitative clinical science. *Clin Pharmacol Ther* 30:564-77, 1981.

Fisher RA: *Statistical Methods and Scientific Inference*. Edinburgh: Oliver and Boyd, 1956.

Fleiss JL: *Statistical Methods for Rates and Proportions*, 2nd ed. New York: Wiley, 1981.

Fletcher RH, Fletcher SW, Wagner EH: *Clinical Epidemiology—The Essentials*. Baltimore: Williams & Wilkins, 1982.

Friedman GD: *Primer of Epidemiology*, 2nd ed. New York: McGraw-Hill, 1980.

Froom J: An international glossary for primary care. *J Family Pract* 13:673-681, 1981.

Garrison FH: *An Introduction to the History of Medicine*, 4th ed. Philadelphia: Saunders, 1929.

2503012458

Greenwood M: *Epidemics and Crowd Diseases*. London: Williams and Norgate, 1933.

Haupt A, Kane TT: *Population Handbook*, 2nd ed. Washington DC: Population Reference Bureau, 1985.

Huggarth J: *Glossary of Health Care Terminology*. Copenhagen: World Health Organization, 1975.

Holland WW (Ed): *Data Handling in Epidemiology*. London: Oxford, 1970.

Holland WW (Ed): *Evaluation of Health Care*. Oxford: Oxford University Press, 1983.

Holland WW, Detels R, Knox G (Eds): *Oxford Textbook of Public Health*, Vol 3. Oxford: Oxford University Press, 1985.

Ibrahim MA: *Epidemiology and Health Policy*. Rockville, MD: Aspen, 1985.

Jammal A, Allard R, Loslier G (Eds): *Dictionnaire d'épidémiologie*. St Hyacinthe, Maloine, Paris: Edisem, 1988.

Jenicek M, Cleroux R: *Epidémiologie*. St Hyacinthe, Québec: Edisem, 1982.

Kahn HA: *An Introduction to Epidemiologic Methods*. New York: Oxford University Press, 1983.

Kelsey JL, Thompson WD, Evans AS: *Methods in Observational Epidemiology*. New York: Oxford University Press, 1986.

Kendall MG, Buckland AA: *A Dictionary of Statistical Terms*, 4th ed. London: Longman, 1982.

Kleinbaum DG, Kupper LL, Morgenstern H: *Epidemiology—Principles and Quantitative Methods*. Belmont: Lifetime Learning Publications, 1982.

Klug WS, Cummings MR: *Concepts of Genetics*. Columbus OH: Merrill, 1986.

Knox EG (Ed): *Epidemiology in Health Care Planning*. London: Oxford University Press, 1979.

Last JM (Ed): *Marx-Yerushalayim Public Health and Preventive Medicine*, 12th ed. Norwalk, CT: Appleton-Century-Crofts, 1986.

Last JM: *Public Health and Human Ecology*. Norwalk CT: Appleton and Lange, 1987.

Lilienfeld AM, Lilienfeld D: *Foundations of Epidemiology*, 2nd ed. New York: Oxford University Press, 1979.

MacMahon B, Pugh TF: *Epidemiology: Principles and Methods*. Boston: Little, Brown, 1970.

Mausner JS, Kramer S: *Epidemiology*, 2nd ed. Philadelphia: Saunders, 1985.

McDowell I, Newell C: *Measuring Health: A Guide to Rating Scales and Questionnaires*. New York: Oxford University Press, 1987.

Meadows AJ, Gordon M, Singleton A: *Dictionary of New Information Technology*. London: Century, 1982.

Meineri CL: *Clinical Trials*. New York: Oxford University Press, 1986.

Miettinen OS: *Theoretical Epidemiology*. New York: Wiley, 1985.

Morris JN: *Uses of Epidemiology*, 3rd ed. London: Churchill Livingstone, 1975.

Morrison AS: *Screening in Chronic Disease*. New York: Oxford University Press, 1985.

Morton NE: *Outline of genetic epidemiology*. New York: Karger, 1982.

Murphy EA: *The Logic of Medicine*. Baltimore: Johns Hopkins University Press, 1976.

Murphy EA: *A Companion to Medical Statistics*. Baltimore: Johns Hopkins University Press, 1985.

Oldham PD: *Measurement in Medicine*. Baltimore: Johns Hopkins University Press, 1976.

Oxford English Dictionary (OED). London: Oxford University Press, 1971.

Pressat R: *Dictionary of Demography*. English Trans. Ed. Christopher Wilson. Oxford: Blackwell, 1985.

Rimm AA, Hartz AJ, Kalbkleisch JH, Anderson AJ, Hoffmann RG: *Basic Biostatistics in Medicine and Epidemiology*. New York: Appleton-Century-Crofts, 1980.

Rothman KJ: *Modern Epidemiology*. Boston: Little, Brown, 1986.

Rothman KJ (Ed): *Causal Inference*. Chestnut Hill, MA: Epidemiology Resources, Inc., 1988.

Rumeau-Rouquette C, Breart G, Padieu R: *Méthodes en épidémiologie*. Paris: Flammarion, 1980.

Schlesselman JJ: *Case-Control Studies; Design, Conduct, Analysis*. New York: Oxford University Press, 1982.

Schuman SH: *Practice-Based Epidemiology*. New York: Gordon and Breach, 1986.

Silverman WA: *Human Experimentation: A Guided Step into the Unknown*. Oxford: Oxford University Press, 1985.

Skinner HA: *The Origin of Medical Terms*, 2nd ed. Baltimore: Williams & Wilkins, 1961.

Snedecor GW, Cochran WH: *Statistical Methods*, 7th Ed. Ames, IO: Iowa University Press, 1979.

Show J: *On the Mode of Communication of Cholera*, 2nd ed. London: Churchill, 1855. [Reprinted 1936 by the Commonwealth Fund of New York, with an Introduction by Wade Hampton Frost; and again by Hafner, New York, 1973.]

Sohni ED (Ed): *Glossary of Evaluation Terms*. Geneva: United Nations, 1978.

Stalleybrass CO: *The Principles of Epidemiology*. London: Routledge, 1931.

Stedman's Medical Dictionary, 22nd ed. Baltimore: Williams & Wilkins, 1972.

Susser MW: *Causal Thinking in the Health Sciences*. New York: Oxford University Press, 1973.

Susser MW, Watson W, Hopper K: *Sociology in Medicine*, 3rd ed. New York: Oxford University Press, 1985.

Swaroop S: *Introduction to Health Statistics*. Edinburgh: Livingstone, 1960.

Tonia B (Ed): *Glossaire d'épidémiologie animale*. Mimeographed. Ecole Nationale Vétérinaire D'Alfort, 1987.

van de Walle E: *Multilingual Demographic Dictionary; English Section*. Liège: Ordina Editions, 1982 (For the International Union for the Scientific Study of Population).

U.S. Department of Health and Human Services Task Force on Health Risk Assessment: *Determining Risks to Health—Federal Policy and Practice*. Dover, MA: Auburn House, 1986.

U.S. House of Representatives: *A Discursive Dictionary of Health Care*. Washington DC: US Government Printing Office, 1976.

Webster's Third New International Dictionary. Unabridged. Springfield, MA: Merriam, 1971.

Weed LL: *Medical Records, Medical Education and Patient Care*. Cleveland: Case Western Reserve University Press, 1969.

White KL, Henderson M (Eds): *Epidemiology as a Fundamental Science*. New York: Oxford University Press, 1976.

Wilson C: *The Dictionary of Demography* (Translation from the French of Roland Pressat). Oxford: Blackwell, 1985.

World Bank: *A Glossary of Population Terminology*. Washington: World Bank, 1985.

World Organization of National Colleges, Academies (WONCA) of Family Practice: *International Classification of Health Problems in Primary Care*, 3rd ed. Oxford: Oxford University Press. In Press, 1987.

6

2503012460

2503012461

News &
Numbers

News & Numbers

A GUIDE TO REPORTING STATISTICAL
CLAIMS AND CONTROVERSIES IN HEALTH
AND OTHER FIELDS

Victor Cohn

SENIOR WRITER AND COLUMNIST, FORMER SCIENCE EDITOR,
Washington Post

FOREWORD BY Frederick Mosteller
ROGER I. LEE PROFESSOR EMERITUS OF MATHEMATICAL STATISTICS,
Harvard University

*A Project of the Center for Health Communication
Harvard School of Public Health*

 IOWA STATE UNIVERSITY PRESS / AMES

2503012462

RA407
C64
1989

A Note to

THE rule
fied. They are
stated or implied
porting, busi
This gu
language of
about the ma
on so
project of the
health and t
ciples and m
used by inqu
a scientific re
environment
weigh and
shows how t

2503012463

© 1989 Victor Cohn. All rights reserved

Composed by Iowa State University Press
Printed in the United States of America

No part of this book may be reproduced in any form or by any electronic or mechanical
means, including information storage and retrieval systems, without written permission
from the publisher, except for brief passages quoted in a review.

First edition, 1989

Library of Congress Cataloging-in-Publication Data

Cohn, Victor, 1919-
News & numbers.

"A project of the Center for Health Communication, Harvard School of Public
Health."
1. Public health—Statistics. 2. Environmental health—Statistics. 3. Vital
statistics. I. Harvard School of Public Health. Center for Health Communi-
cation. II. Title. III. Title: News and numbers.
RA407.C64 1989 362.1'021 88-6807
ISBN 0-8138-1442-1
ISBN 0-8138-1457-5 (pbk.)

2503012464

A Note to Readers

THE rules of statistics are the rules of good thinking, codified. They apply to any kind of reporting in which numbers—stated or implied—are involved: political reporting, science reporting, business, economics, sports, or whatever.

This guide is an attempt to explain the role, logic, and language of statistics, so we reporters can ask better questions about the many alleged facts or findings that rest, or should rest, on some credible numbers. Because this manual began as a project of the Harvard School of Public Health, the reporting of health and the environment is the major example. But the principles and many of the suggested "questions for reporters" can be used by inquiring reporters in any field. They can help you read a scientific report or listen to the conflicting claims of politicians, environmentalists, physicians, scientists, or almost anyone and weigh and explain them. And the final chapter specifically shows how these principles apply in all areas.

VICTOR COHN

Contents

FOREWORD BY Frederick Mosteller, ix

ACKNOWLEDGMENTS, xi

1. Facts and Figures—We Can Do Better, 3
2. The Certainty of Uncertainty, 8
3. The Scientific Way, 12
Probability, 14
"Power" and Numbers, 20
Bias and Confounders, 24
Variability, 30
4. Studies, Good and Bad, 35
Experiments versus Seductive Anecdotes, 37
Clinical Trials, 38
What Makes a Study Honest? 40
Epidemiology: Hippocrates to AIDS, 43
5. Questions Reporters Can Ask, 48
6. Tests and Testing, 64
Drugs and Drug Trials, 68
Animals as Models for Us, 72
7. Vital Statistics: The Numbers of Life and Health, 74
Crude Rates versus Rates That Compare, 76
Other Ways to Compare, 78
Reporting Hospital Death Rates, 79
Cancer Rates and Cancer "Cures", 86
The Important Questions about Cancer, 88
Shifts, Drifts, and Blips, 96

vii

2503012465

8. The Statistics of Environment and Risk, 98

- Who's Believable? 107
- Questions to Ask, 108
- Evaluating Environmental Hazards, 116
- Advice from Reporters, 121

9. The Statistics of Politics, Economics, and Democracy, 126

- The State of the Nation's Statistics, 146
- The Bottom Line, 151

WHERE TO LEARN MORE: *A Bibliography and Other Sources*, 153**NOTES, 157****GLOSSARY/INDEX, 165****Foreword**

REPORT science to the accuracy. Although stories, the bi presents spec misleading misstent," and "sults that are laymen definitio siderable diffe

Science v such as biosta have been imp ertheless, they permanent fo

Victor C ual to help all wants to give facts or mysti

Cohn's bx Science Policy Research and that faculty n have been abl

2503012466

Foreword

REPORTERS play an essential role in communicating science to the public. In common with scientists, they desire accuracy. Although health and medicine provide many exciting stories, the biostatistics that scientists must use in their studies presents special problems for reporters. It gives uncommon and misleading meanings to common words like "significant," "consistent," and "power." Mathematical statistics often produces results that are disturbingly counterintuitive, at least at first, to laymen and scientists alike. In vital statistics and epidemiology, definitions often seem arbitrary, and slight changes make considerable differences in the findings.

Science writers often take short courses in special topics such as biostatistics. I have taught in some of these courses and have been impressed by the seriousness of the participants. Nevertheless, they need some of this material in an accessible and permanent form.

Victor Cohn of the *Washington Post* has prepared this manual to help all reporters cut through these statistical tangles. He wants to give them a guide to the ways that statistics can clarify facts or mystify the reader.

Cohn's book grew out of the Media Project of our Health Science Policy Working Group of the Division of Health Policy Research and Education at Harvard University. I am pleased that faculty members of the Harvard School of Public Health have been able to help him produce this book as a visiting fellow.

FOREWORD

x
in 1978 and 1984 and as a contributor to the Health Science Policy Working Group.

Through the Media Project, with the help of Jay Winsten, we have also examined sources of pressures on the science writer.¹ In the future we want to use what we have learned through many discussions with science writers to advise scientists on their role in the media.

By such efforts, including this book, and by many similar efforts in this and other fields, scientists and writers may gradually upgrade the whole communication system, scientific and journalistic. Thus we may clear the communication channel between science and the public.

FREDERICK MOSTELLER

Acknowl

MY mai
has been Dr.
tus of mathe
partments of
Harvard Sch
energy, and
for the fact
approach rat
stater

T.
tion, and by
ing, which p
journalists, i

I did m
School of Pu
Center for I
guide, and
Thomas A.
were Drs. C
Kaiser organ
and Peter N
writings I le
Cass Petersc
Runkle, no

I also c

2503012468

Acknowledgments

My main mentor and guide in the preparation of this book has been Dr. Frederick Mosteller, Roger I. Lee professor emeritus of mathematical statistics and former chairman of the departments of Biostatistics and Health Policy and Management, Harvard School of Public Health. He gave so fully of his time, energy, and knowledge that he should be listed as coauthor but for the fact that I sometimes used a journalist's freewheeling approach rather than a statistician's rigor. This makes any misstatements mine.

The project was supported by the Russell Sage Foundation, and by the Council for the Advancement of Science Writing, which pointed the way in holding seminars on statistics for journalists, including the first of its kind in 1964.

I did much of the work as a visiting fellow at the Harvard School of Public Health, where Dr. Jay Winstein, director of the Center for Health Communication, was another indispensable guide, and Drs. John Bailar III, Nan Laird, Philip Lavin, Thomas A. Louis, and Marvin Zelen were valuable helpers. As were Drs. Gary D. Friedman and Thomas M. Vogt of the Kaiser organizations, Michael Greenberg of Rutgers University, and Peter Montague of Princeton University (on all of whose writings I leaned); Lewis Cope of the Minneapolis *Star Tribune*; Cass Peterson of the *Washington Post*, and my daughter, Deborah Runkle, no mean statistician.

I also owe thanks to Harvard's Drs. Peter Braun, Harvey

ACKNOWLEDGMENTS

Fineberg, Howard Frazier, Howard Hiatt, William Hsaio, Herb Sherman, and William Stason. And to Drs. Stuart A. Bessler, Syntex Corporation; H. Jack Geiger, City University of New York; Nicole Schupf Geiger, Manhattanville College; Charles Moertel, Mayo Clinic; Arnold Relman, *New England Journal of Medicine*; Eugene Robin, Stanford University; and Sidney Wolfe, Public Citizen Health Research Group. Also Katherine Wallman, Council of Professional Associations on Federal Statistics; Howard L. Lewis, American Heart Association; Philip Meyer, University of North Carolina; Mildred Spencer Sanes; Earl Ubell, WCBS-TV, New York City; and Philip Hilts, Cristine Russell, and Barry Sussman, *Washington Post*. I am indebted to my editors at the *Washington Post*, particularly Abigail Trafford, Ben Cason, Carol Kruoff, Len Downie, and Howard Simons for their understanding and support.

The work was also aided by the Andrew W. Mellon Foundation. The American Cancer Society, American Heart Association, Commonwealth Fund, Gannett Foundation, Henry J. Kaiser Family Foundation, Mayo Medical Resources, Milbank Memorial Fund, Pew Charitable Trusts, Philip L. Graham Fund, Russell Sage Foundation, and John Cowles, Jr., have contributed to this manual's initial distribution.

2503012470

2503012471

News &
Numbers

Facts and Figures — We Can Do Better

Facts and Figures! Put 'em Down!

—Charles Dickens (in *The Chimes*)

There are lies, there are damned lies, and there are statistics.

—Disraeli

Almost everyone has heard that "figures don't lie, but liars can figure." We need statistics, but liars give them a bad name, so to be able to tell the liars from the statisticians is crucial.

—Dr. Robert Hooke

We journalists like to think we deal mainly in facts and ideas, but much of what we report is based on numbers. Politics comes down to votes. Budgets and dollars dominate government. The economy, business, employment, sports—all demand numbers.

The environment, pollutants, toxic chemicals. Again, we see counts and measurements and, most likely, widely varying estimates, some careful, some questionably high or low. An environmentalist says a nuclear power plant or toxic waste dump will cause so many cases of cancer. An industry spokesman denies it. What are their numbers? Where did they get them? How valid are they?

A doctor reports a promising, even exciting new treatment. Is the claim justified or based on a biased or unrepresentative sample? Or too few patients to justify any claim? Science, medicine, technology, the weather, intelligence—all are statistical.

2503012472

3

CHAPTER 1

FACTS AND FIG

Science is observation, experimentation, measurement, and all these involve numbers, whether we reporters pay attention to them or not.

Statistics are used or misused even by people who tell us, "I don't believe in statistics," then claim that all of us or most people or many do such and such. The question for reporters is, how should we not merely repeat such numbers, stated or implied, but also interpret them to deliver the best possible picture of reality?

We can be better reporters if we understand how the best statisticians—the best figurers—figure. And if we learn a few questions to help us separate the wheat from the chaff.

I do not say that telling the truth—describing reality—will then become easy, for we are constantly bombarded with sweeping claims in convincing wrappings, and the disputed subjects are endless. Medical and surgical treatments, radiation, pesticides, nuclear power, the probability of environmental disasters, the side effects of medicines—almost nothing seems settled.

Like it or not, we must wade in. Whether we will it or not, we have in effect become part of the regulatory apparatus. Dr. Peter Montague of Princeton University tells us, "The environmental and toxic situation is so complex, we can't possibly have enough officials to monitor it. Reporters help officials decide where to focus their activity."

"Journalists opened up" the Love Canal toxic waste issue by "independent investigation," according to Cornell University's Dr. Dorothy Nelkin. "The extensive press coverage contributed to investigations that eventually forced the re-staffing of the Environmental Protection Agency and the creation of a national toxic waste disposal program."¹

That very coverage, however, may also have stampeded public officials into hasty, ill-conceived studies that left unanswered the crucial question: Did the Love Canal wastes actually cause birth defects and other physical problems?² The very way we report a medical or environmental controversy can affect the outcome. If we ignore a bad situation, the public may

suffer. If we v
"no danger," t
experimental
false hope.

It is not
National Car
refuse to cor
think "carcin
persons prob
cancers are ei
most informe
related mainl
and very pos
percent of al
carcinogens—
foods, air, an

When it
issues, or wi
making the s
state or unde
of Pr
statist.
terpretation.
evident; you
negative]. A
sterile is mor
that apple p

We also
the space or
news directo
story yet." E
done. In a
major south
traction afte
who worked
numbers fix

2503012473

suffer. If we write "danger," the public may quake. If we write "no danger," the public may be falsely reassured. If we paint an experimental medical treatment too brightly, the public is given false hope.

It is not just what we write, it is what we emphasize. A National Cancer Institute survey indicated that many persons refuse to consider healthy changes in life-style because they think "carcinogens are everywhere in the environment." Such persons probably have read or heard again and again that most cancers are environmentally related, although, in the opinion of most informed scientists, most fatal "environmental" cancers are related mainly to individual behavior, outstandingly smoking, and very possibly diet. By various estimates, perhaps 5 to 15 percent of all cancers are related to exposures to man-made carcinogens—chemicals we have inserted into the workplace, foods, air, and water.³

When it comes to such emotionally charged and complex issues, or when it simply comes to running for page one or making the six o'clock news, the best among us sometimes overstate or understate. Philip Meyer, veteran reporter and author of *Precision Journalism*, writes, "Journalists who misinterpret statistical data usually tend to err in the direction of overinterpretation. . . . The reason for this professional bias is self-evident; you usually can't write a snappy lead upholding [the negative]. A story purporting to show that apple pie makes you sterile is more interesting than one that says there is no evidence that apple pie changes your life."⁴

We also work fast, sometimes too fast, with severe limits on the space or time we may fill. We find it hard to tell editors or news directors, "I haven't had enough time. I don't have the story yet." Even a long-term project or special may be hurriedly done. In a newsroom "long-term" may mean a few weeks. A major southern newspaper had to print a long, front-page retraction after a series of front-page stories alleged that people who worked at or lived near a plutonium plant suffered in excess numbers from a blood disease. "Our reporters obviously had

2503012474

confused statistics and scientific data," the editor admitted. "We did not ask enough questions."

We tend to oversimplify. We may report, "A study showed that black is white" or "So-and-so announced that . . .," when a study merely suggested that there was some evidence that such might be the case. We may slight or omit the fact that a scientist calls a result "preliminary." As scientific unsophisticates, we may confuse a study that merely suggests a hypothesis that should be investigated—very frequently the case—with a study that presents strong and conclusive evidence.

We often omit essential perspective, context, or background. Dr. Thomas Vogt of the Kaiser Permanente Center for Health Research tells of seeing the headline "Heart Attacks From Lack of C" and then, two months later, "People Who Take Vitamin C Increase Their Chances of a Heart Attack." Both stories were based on limited, and far from conclusive, animal studies.

Scientists who do poor studies or overstate their results deserve part of the blame. But bad science is no excuse for bad journalism. We tend to rely most on "authorities" who are either most quotable or quickly available or both, and they often tend to be those who get most carried away with their sketchy and unconfirmed but "exciting" data—or have big axes to grind, however lofty their motives. The cautious, unbiased scientist who says, "Our results are inconclusive" or "We don't have enough data yet to make any strong statement" or "I don't know" tends to be omitted or buried someplace down in the story.

We are influenced too by intense and growing competition to tell the story first and tell it most dramatically. I was once asked by a Harvard researcher, "Does competition affect the way you present a story?" I thought and had to answer, "We have to almost overstate. We have to come as close as we can within the boundaries of truth to a dramatic, compelling statement. A weak statement will go no place." Another reporter said, "The fact is, you are going for the strong [lead and story]. And, while

not patently a year later."

We report fear. A new "new alarm is

Alarms : maxim: Bad Jack Geiger, a sor of medici

I know I wrote wrongly. I wrote written. I wrote clear later that hadn't asked— capable of getting say something factors?"

H

2503012475

FACTS AND FIGURES: WE CAN DO BETTER

7

not patently absurd, it may not be the lead you would go for a year later.⁷

We reporters are also subject to human hope and human fear. A new "cure" comes along, and we want to believe it. A new alarm is sounded, and we too tremble.

Alarms also make news. We too often obey a sardonic maxim: Bad news is good news; good news is no news. Dr. H. Jack Geiger, a respected former science writer and now a professor of medicine, says,

I know I wrote stories in which I explained or interpreted the results wrongly. I wrote stories that didn't have the disclaimers I should have written. I wrote stories under competitive pressure, when it became clear later that I shouldn't have written them. I wrote stories when I hadn't asked — because I didn't know enough to ask — Was your study capable of getting the answers you wanted? Could it be interpreted to say something else? Did you take into account possible confounding factors?⁸

How can we learn to do better?

2503012476

The Certainty of Uncertainty

2

Too much of the science reporting in the press [blurs] what we're sure of and what we're not very sure of and what is inconclusive. The notion of tentativeness tends to drop out of much reporting.

—Dr. Harvey Brooks

The only trouble with a sure thing is the uncertainty.

—Author unknown

THE first thing to understand about science is that it is almost always uncertain. A scientist, seeking to explain or understand something—be it the behavior of an atom or the effect of the toxic chemicals at a Love Canal—usually proposes a hypothesis, then seeks to test it by experiment or observation. If the evidence is strongly supportive, the hypothesis may then become a theory or at some point even a law, like the law of gravity.

A theory may be so solid that it is generally accepted. Example: the theory that cigarette smoking causes lung cancer, for which almost any reasonable person would say the case has been proved, for all practical purposes. The phrase "for all practical purposes" is important, for scientists, being practical people, must often speak at two levels: the strictly scientific level and the level of ordinary reason that we require for daily guidance.

Example: In June 1985, 16 forensic experts examined the bones that were supposedly those of the "Angel of Death," Dr. Josef Mengele. Dr. Lowell Levine, delegated by the Department of Justice, then said, "The skeleton is that of Josef

8

Mengele with: cos Segre of t with the law c cians." Pushed important mat of the patholog findings.¹ (Lat

But all at tainty in almos bility that suc

Widely b wholly or part say," reports I Journal of Medi help the public with an eleme a probable na not certainty. V best opinion a future."

Exa. mended that v cal cancer. Th three years fo Statistics had matter is still changed agair

Scientists a failing. Whe sionally shows right, the scie ing.

The publ have a hard t sions. We all today and anc show discussi

2503012477

2503012478

THE CERTAINTY OF UNCERTAINTY

9

Mengele within a reasonable scientific certainty" and Dr. Marcos Segre of the University of Sao Paulo, explained, "We deal with the law of probabilities. We are scientists and not magicians." Pushed by reporters' questions — after all, this was an important matter, and what should the public believe? — several of the pathologists said they had "absolutely no doubt" of their findings.¹ (Later evidence made the case even stronger.)

But all any scientist can scientifically say — say with certainty in almost any such case — is, there is a very strong probability that such and such is true.

Widely believed theories or conclusions are often proved wholly or partly wrong. "When it comes to almost anything we say," reports Dr. Arnold Relman, editor of the *New England Journal of Medicine*, "you, the reporter, must realize — and must help the public understand — that we are almost always dealing with an element of uncertainty. Most scientific information is of a probable nature, and we are only talking about probabilities, not certainty. What we are concluding is the best we can do, our best opinion at the moment, and things may be updated in the future."

Example: Until 1980 the American Cancer Society recommended that women have an annual Pap smear to detect cervical cancer. The recommendation was then changed to every three years for many women, after two initial examinations. Statistics had shown that this would be equally effective.² The matter is still controversial, and the recommendation has been changed again in the light of new knowledge.

Scientists are often wrong. In science this is not necessarily a failing. When new evidence disproves an old theory, or occasionally shows that some little believed, even kooky notion is right, the scientific method is doing what it should. It is working.

The public, and even some reporters and especially editors, have a hard time understanding these sometimes drastic revisions. We all hear the question, Why do they say one thing today and another thing tomorrow? I was once on a radio talk show discussing unsettled medical controversies when a testy

listener phoned in to exclaim, "They say' is a damned liar!"

"They" of course may be different theys who arrive at different conclusions about inconclusive evidence in a thousand areas: the role of fats and cholesterol in the diet, the effects of low-level radioactivity, the cause of the extinction of dinosaurs.

Why so much uncertainty? Science is always a continuing story. Nature is complex, and almost all methods of observation and experiment are imperfect. "There are flaws in all studies," says Harvard's Dr. Marvin Zelen.³ There may be weaknesses, often unavoidable ones, in the way a study is designed or conducted. Observers are subject to human bias and error. Subjects fluctuate. Measurements fluctuate.

Many studies are thus inconclusive, and virtually no single study proves anything. "Fundamentally," writes Dr. Thomas Vogt, "all scientific investigations require confirmation, and until it is forthcoming all results, no matter how sound they may seem, are preliminary."⁴

Medicine, in particular, is full of disagreement and controversy. "No clinical trial is ever perfect," Harvard's Dr. John Bailar observes. Unlike new drugs, medical treatments and tests and surgical operations need not even be subjected to experimental studies before being applied. "Most treatments escape and will continue to escape rigorous evaluation," Bailar says.⁵

The reasons are many: lack of funds to mount enough trials; lack of enough patients at any one center to mount a meaningful trial; the expense and difficulty of doing multicenter trials; the swift evolution and obsolescence of medical techniques; the fact that, with the best of intentions, medical data—histories, physical examinations, interpretations of tests, descriptions of symptoms and diseases—are notoriously inexact and vary from physician to physician; and the serious ethical obstacles to trying a new procedure when an old one is doing some good, or to experimenting on children, pregnant women, or the mentally ill.

While all studies have flaws, some have more flaws than others. Study after study has found that many articles in the most prestigious medical journals are replete with shaky statis-

tics and lack of patients' complications. Papers previously reported by the mere progressive results than or criticism of uncertain findings.

The upshot of the organization's care is based on . . . Seemingly doctrinaire, perhaps out to be supported be found."

In general, possible benefit that only a few cancer. Only less dramatic economy, or is rich in area or statistically discarded.

Occasionally, results. More often data that control method ascribing fraud in mind the competence to

So some tainty need to survive on the policy, to govern basis of incor can do so.

2503012479

tics and lack of any explanation of such crucial matters as patients' complications and the number of patients lost to follow-up. Papers presented at medical meetings, many of them widely reported by the media, are even less reliable. Many papers are mere progress reports on incomplete studies. Some state tentative results that later collapse. Some are given to draw comment or criticism or get others interested in a provocative but still uncertain finding.⁶

The upshot, according to Dr. Gary Friedman of the Kaiser organization's Permanente Medical Group: "Much of health care is based on tenuous evidence and incomplete knowledge. . . . Seemingly authoritative statements and accepted medical doctrines, perpetuated through textbook and lectures, often turn out to be supported by the most meager of evidence, if any can be found."⁷

In general, possible risks tend to be underestimated and possible benefits overestimated. For decades surgeons swore that only a radical mastectomy was the treatment for breast cancer. Only recently were clinical trials mounted to show that less drastic treatments seem equally effective. Prefrontal lobotomy, overstrict bed rest, drugs by the carload—medical history is rich in treatments that were given for years without question or statistically rigorous study, only to be proved wrong and discarded.

Occasionally, unscrupulous investigators falsify their results. More often, they may wittingly or unwittingly play down data that contradict their theories, or they may search out statistical methods that give them the results they want. Before ascribing fraud, says Harvard's Dr. Frederick Mosteller, "keep in mind the old saying that most institutions have enough incompetence to explain almost any results."⁸

So some uncertainty almost always prevails. But uncertainty need not stand in the way of good sense. To live—to survive on this globe, to maintain our health, to set public policy, to govern ourselves—we almost always must act on the basis of incomplete or uncertain information. There is a way we can do so.

2503012480

The Scientific Way 3

Somehow the wondrous promise of the earth is that there are things beautiful in it, things wondrous and alluring, and by virtue of your trade, you want to understand them.

— Mitchell Feigenbaum
Cornell University physicist and mathematician

The great tragedy of Science—the slaying of a beautiful hypothesis by an ugly fact.

— Thomas Henry Huxley

TO reporters, the world is full of true believers, peddling their "truths." The sincerely misguided and the outright fakers are often highly convincing, also newsy. How can we tell the facts, or the probable facts, from the chaff?

We can borrow from science. We can try to judge all possible claims of fact by the same methods and rules of evidence that scientists use to derive some reasonable guidance in scores of unsettled issues.

As a start, we can ask these questions:

How do you know?

Have the claims been subjected to any studies or experiments?

Were the studies acceptable ones, by general agreement? For example: Were they without any substantial bias?

Have results been fairly consistent from study to study?

Have the findings resulted in a consensus among others in the same field? Do at least the majority of informed persons agree? Or should we withhold judgment until there is more evidence?

Always: Are the conclusions backed by believable statistical evidence?

And what is the c
be?

Obviously,
rather than nur
that reporters c

There are
useful ones: T.
interpreting da
a way of extra
of mathematic

Statistics c
and inexpert s
be difficult for
possible. Once
in almost all.

There are
"Edison had i
author. "It doe
It did not take
ton's et al. an
centur
until then ha

Overwhe
probability, i
and the use c
called the or
many events
women, yet i
before it bec
develop hea
some years
was to wom

The be
line (for exa
a study is a

2503012481

And what is the degree of certainty or uncertainty? How sure can you be?

Obviously, much of statistics involves attitude or policy rather than numbers. And much, at least much of the statistics that reporters can most readily apply, is good sense.

There are many definitions of statistics as a tool. A few useful ones: The science and art of gathering, analyzing, and interpreting data; a means of deciding whether an effect is real; a way of extracting information from a mass of raw data; a set of mathematical processes derived from probability theory.

Statistics can be manipulated by charlatans, self-deluders, and inexpert statisticians. Deciding on the truth of a matter can be difficult for the best statisticians, and sometimes no decision is possible. Uncertainty will ever rule in some situations and lurk in almost all.

There are rare situations in which no statistics are needed. "Edison had it easy," says Dr. Robert Hooke, a statistician and author. "It doesn't take statistics to see that a light has come on."¹ It did not take statistics to tell 19th-century physicians that Morton's ether anesthesia permitted painless surgery or to tell 20th-century physicians that the first antibiotics cured infections that until then had been highly fatal.

Overwhelmingly, however, the use of statistics, based on probability, is called the soundest method of decision making, and the use of large numbers of cases, statistically analyzed, is called the only means for determining the unknown cause of many events. Birth control pills were tested on several hundred women, yet the pills had to be used for several years by millions before it became unequivocally clear that some women would develop heart attacks or strokes. The pills had to be used for some years more before it became clear that the greatest risk was to women who smoked and women over 35.

The best statisticians, let alone practitioners on the firing line (for example, physicians), often have trouble deciding when a study is adequate or meaningful. Most of us cannot become

2503012482

statisticians, but we can at least learn that there are studies and studies, and the unadorned claim "We made a study" or "We did an experiment" may not mean much. We can learn to ask more pointed questions if we understand some basic concepts and other facts about scientific studies.

These are some bedrock statistical concepts:

- Probability
- "Power" and numbers
- Bias and confounders
- Variability

Probability

Scientists cope with uncertainty by measuring probabilities. Since all experimental results and all events can be influenced by chance and almost nothing is 100 percent certain in science and medicine and life, probabilities sensibly describe what has happened and should happen in the future under similar conditions. Aristotle said, "The probable is what usually happens," but he might have added that the improbable happens more often than most of us realize.

The accepted numerical expression of probability in evaluating scientific and medical studies is the *P* (or *probability*) value. The *P* value is one of the most important figures a reporter should look for. It is determined by a statistical formula that takes into account the numbers of subjects or events being compared in order to answer the question, could a difference or result this great or greater have occurred *by chance alone*? By more precise definition, the *P* value expresses the probability that an observed relationship or effect or result could have *seemed* to occur by chance *if there had actually been no real effect*. A low *P* value means a low probability that this happened, that a medical treatment, for example, might have been declared beneficial when in truth it was not.

Here is why the *P* value is used to evaluate results. A

scientific inves
commonly sets
hypothesis: that
back the origi
pothesis. The
number or as
"greater than"
pened, that th
chance—or, to
variation.

• By con
only 5 or few
pened by cha
usually called
ues are used).
ally implies th

• *A higher*
statistically sig
result is due

Ir
the sho
dinary logic.
replaces "it c

Why the
People have
purposes. Ar
Mosteller tel
class and aft
cious going c
the fifth hea
chance in 16
that five hea
there is some
neighborho

Another
late a confide

2503012483

scientific investigator first forms a hypothesis. Then he or she commonly sets out to try to disprove it by what is called the *null hypothesis*: that there is no effect, that nothing will happen. To back the original hypothesis, the results must reject the null hypothesis. The *P* value, then, is expressed either as an exact number or as $< .05$, say, or $> .05$, meaning "less than" or "greater than" a 5 percent probability that nothing has happened, that the observed result could have happened just by chance—or, to use a more elegant statistician's phrase, by *random variation*.

- By convention, a *P* value of $.05$ or less, meaning there are only 5 or fewer chances in 100 that the result could have happened by chance, is most often regarded as *low*. This value is usually called *statistically significant* (though sometimes other values are used). The unadorned term "statistically significant" usually implies that *P* is $.05$ or less.
- A higher *P* value, one greater than $.05$, is usually seen as not statistically significant. The higher the value, the more likely the result is due to chance.

In common language, a low chance of chance alone calling the shots replaces the "it's certain" or "close to certain" of ordinary logic. A strong chance that chance could have ruled replaces "it can't be" or "almost certainly can't be."

Why the number $.05$ or less? Partly for standardization. People have agreed that this is a good cutoff point for most purposes. And partly out of old friend common sense. Frederick Mosteller tells us that if you toss a coin repeatedly in a college class and after each toss ask the class if there is anything suspicious going on, "hands suddenly go up all over the room" after the fifth head or tail in a row. There happens to be only 1 chance in $16 - .0625$, not far from $.05$, or 5 chances in $100 - 1$ that five heads or tails in a row will show up in five tosses, "so there is some empirical evidence that the rarity of events in the neighborhood of $.05$ begins to set people's teeth on edge."¹²

Another common way of reporting probability is to calculate a *confidence level*, as well as a *confidence interval* (or *confidence*

2503012484

limits or range). This is what happens when a political pollster reports that candidate X would now get 50 percent of the vote and thereby lead candidate Y by 3 percentage points, "with a 3-percentage-point margin of error plus or minus and a 95 percent confidence level." In other words, Mr. or Ms. Pollster is 95 percent confident that X's share of the vote would be someplace between 53 and 47 percent. Similarly, candidate Y's share might be 3 percentage points greater (or less) than the figure predicted. In a close election, that margin of error could obviously turn a predicted defeat into victory. And that sometimes happens.

An important point in looking at the results of political polls (and any other statements of confidence): In the reports we read, the plus or minus 3 (or whatever) percentage points is often omitted, and the pollster merely mentions a "3-point margin of error." This means there is actually a 6-point range within which the truth *probably* lurks.

The more people who are questioned in a political poll or the larger the number of subjects in a medical study, the greater the chance of a high confidence level and a narrow, and therefore more reassuring, confidence interval.

No matter how reassuring they sound, *P* values and confidence statements cannot be taken as gospel, for .05 is not a guarantee, just a number. There are several important reasons for this.

- All that *P* values measure is the *probability* that the results might have been produced by some sneaky random process. In 20 results where only chance is at work, 1, on the average, will have a reassuring-sounding but misleading *P* value of <.05. One, in short, may be a false positive.

Dr. Marvin Zelen points out that there may be 6,000 to 10,000 clinical (medical) trials of cancer treatment under way today, and if the conventional value of .05 is adopted as the upper permissible limit for false positives, then every 100 studies with no actual benefit may, on average, produce 5 false-positive results. Hence, we may expect 50 false positive results, on

average, for ever fact has said, "With chemotherapy ir therapies in the paths.

Amazingly, tested. Scientist negative results them. Nor are sing studies that firmatory studie

- Statistical cause and effec member the ro Unless an assoc that the case is ing more study

To statistic ference betwee there is correlation

- If the nu value may sin detect somethijects. Highly "sble differences

- An imp other variable not taken into

- Statistic cal—that is, n rienced report and jump to t called their st tween two lar

2503012485

average, for every 1,000 trials with no beneficial effects! Zelen in fact has said, "We may now have reached an impasse in cancer chemotherapy in which there are large numbers of false-positive therapies in the clinic," leading physicians down many false paths.

Amazingly, most false positives probably remain undetected. Scientists do not profit much professionally by reporting negative results. Journal editors are not keen on publishing them. Nor are scientists keen on doing costly and time-consuming studies that merely confirm someone else's work, so "confirmatory studies are rare," Zelen reports.

- Statistical significance alone does not mean there is a cause and effect. *Correlation* or *association* is not causation. Remember the rooster who thought his crowing made the sun rise? Unless an association is so powerful and so constantly repeated that the case is overwhelming, association is only a clue, meaning more study or confirmation is needed.

To statisticians, incidentally, there is this important difference between correlation and association: *Association* means there is at least a possible relation between two variables. A *correlation* is a measure of the association.

- If the number of subjects is too small, an unimpressive *P* value may simply mean that there were too few subjects to detect something that might have shown an effect in more subjects. Highly "significant" *P* values can sometimes adorn negligible differences in large samples.

- An impressive *P* value might also be explained by some other variable or variables — other conditions or associations — not taken into account.

- Statistical significance does not mean biological, clinical—that is, medical—or practical significance, though inexperienced reporters sometimes see or hear the word "significant" and jump to that conclusion, even reporting that the scientists called their study "significant." Example: A tiny difference between two large groups in mean hemoglobin concentration, or

2503012486

red blood count (say, 0.1 g/100 mL, or a tenth of a gram per 100 milliliters), may be statistically significant yet medically meaningless.⁴

- Eager scientists can consciously or unconsciously manipulate the *P* value by failing to adjust for other factors, by choosing to compare different end points in a study (say, condition on leaving the hospital rather than length of survival), or by choosing the way the *P* value is calculated or reported.

There are several mathematical paths to a *P* value, such as the chi-square (χ^2), *t*, *F*, *z*, and paired *t* tests. All may be legitimate. But be warned. Dr. David Salsburg of Pfizer, Inc., has written in the *American Statistician* of the unscrupulous practitioner who "engages in a ritual known as 'hunting for *P* values'" and finds ways to modify the original data to "produce a rich collection of small *P* values" even if those that result from simply comparing two treatments "never reach the magical .05."⁵

"If you look hard enough through your data," contributes an investigator at a major medical center, "if you do enough subset analyses, if you go through 20 subsets, you can find one"—say, "the effect of chemotherapy on premenopausal women with two to five lymph nodes"—"with a *P* value less than .05. And people do this."

"Statistical tests provide a basis for probability statements," writes Dr. John Bailar, "only when the hypothesis is fully developed before the data are examined. . . . If even the briefest glance at a study's results moves the investigator to consider a hypothesis not formulated before the study was started, that glance destroys the probability value of the evidence at hand." (At the same time, Bailar adds, "review of data for unexpected clues . . . can be an immensely fruitful source of ideas" for new hypotheses "that can be tested in the correct way." And occasionally "findings may be so striking that independent confirmation . . . is superfluous.")⁶

A rather sophisticated—and possibly touchy—line of questioning that some reporters might want to try if they're skeptical: *How did you arrive at your P value? Did you use the test planned in*

advance in your pro:
report the best-soun
And you m.

The laws of
even impossible
We've all tal
and bumped int
don't know, but
work, the chanc
1,024. Yet I wo
year period. W
statisticians call
few people wit
cover, there wil
birth defects th
in a great while

In a large
unusual. They
and ofte
duce uni
evidence. The
large number c
occurred. They
ity are wrong."

"We [rep
dence," Philip
and we are rig
mind our reac
from a few in
member. The

A statistic
people or, a st:
whom such ar
The chance of
oping leukemi

2503012487

advance in your protocol or study design, or did you apply several tests, then report the best-sounding one?

And you may think of other questions.

The laws of probability also teach us to expect some unusual, even impossible-sounding events.

We've all taken a trip to New York or London or someplace and bumped into someone from home. The chance of that? I don't know, but if you and I tossed for a drink every day after work, the chance that I would ever win 10 times in a row is 1 in 1,024. Yet I would probably do so sometime in a four- or five-year period. What I like to call the Law of Unusual Events — statisticians call it the Law of Small Probabilities — tells us that a few people with apparently fatal illnesses will inexplicably recover; there will be some amazing clusters of cases of cancer or birth defects that will have no common cause, and I may once in a great while bump into a friend far from home.

In a large enough population such coincidences are not unusual. They are the rule. They produce striking anecdotes and often striking news stories. In the medical world they produce unreliable, though often cited, testimonial or anecdotal evidence. "The world is large," Vogt notes, "and one can find a large number of people to whom the most bizarre events have occurred. They all have personal explanations. The vast majority are wrong."

"We [reporters] are overly susceptible to anecdotal evidence," Philip Meyer writes. "Anecdotes make good reading, and we are right to use them. . . . But we often forget to remind our readers — and ourselves — of the folly of generalizing from a few interesting cases. . . . The statistic 'is hard to remember. The success stories are not.'

A statistic to ask about is the *denominator* — the number of people or, a statistician would say, the *population or domain* — in whom such an event might happen. Zelen cites this example: The chance of any youngster between ages five and nine developing leukemia is 3 in 100,000 per year. In a school with 100

2503012488

children of this age group, we would expect only 3 cases in 100 years. But in this nation with thousands of schools, we would occasionally—such is chance—find schools with 3 or more cases in a single year. "Then one is faced with the problem of interpretation," Zelen says. "Is this one of those rare events that is surely going to be observed? Or is it due to some causal factor?"

A reporter in this instance might ask a statistician at the National Cancer Institute or a medical center, What is the chance of such an event in such a population? How many similar unusual events are probably never reported?

"Power" and Numbers

This gets us to another statistical concept: *power*. Statistically, "power" means the probability of finding something if it's there. Example: Given that there is a true effect, say a difference between two medical treatments or an increase in cancer caused by a toxin in a group of workers, how likely are we to find it?

Sample size confers power. Statisticians say, "Funny things can happen in small samples without meaning very much" . . . "There is no probability until the sample size is there" . . . "Large numbers confer power" . . . "Large numbers at least make us sit up and take notice."*

All this concern about sample size can also be expressed as the *law of large numbers*, which says that as the number of cases increases, the probable truth of a conclusion or forecast increases. The *validity* (truth or accuracy) and *reliability* (reproducibility) of the statistics begin to converge on the truth.

We already learned this when we talked about probability.

*There is another unrelated use of the word "power." Scientists commonly speak of increasing or "raising" some quantity by a power of 2 or 3 or 100 or whatever. "Power" here means the product you get when you multiply a number by itself one or more times. Thus, in $2 \times 2 = 4$, 4 is the second power of 2, or to put it another way, there are two 2's in your equation. This is commonly written 2^2 and known as 2 to the second power or just 2 to the second. In $2 \times 2 \times 2 = 8$, 8 has been raised to the third power. When you think about 2^{100} , you see the need for the shorthand.

But by thinkin
both sample si:
too affects the p
if the number
shift from succe
cally decrease

If six pati
rate, the shift
success rate to
any case that i
valid or accur
not have reliab
samples. The
no fatal biases
would have ir

One can:

I have m:
claim, I' k
finding
example, som
Would it seem

Or if the
100 percent ir
total and subc
changed, exce
analysis. But
times try thre
problem or er

A statis
not large on
counts of ar
has 10,000
thing new !

2503012489

THE SCIENTIFIC WAY

21

But by thinking of power as statisticians do—as a function of both sample size and the accuracy of measurement, since that too affects the probability of finding something—we can see that if the number of treated patients is small in a medical study, a shift from success to failure in only a few patients could dramatically decrease the success rate.

If six patients have been treated with a 50 percent success rate, the shift to the failure column of just one would cut the success rate to 33 percent. And the total number is so small in any case that the result has little reliability. The result might be valid or accurate, but it would not be generalizable—it would not have reliability until confirmed by careful studies in larger samples. The larger the sample, and assuming there have been no fatal biases or other flaws, the more confidence a statistician would have in the result.

One canny science reporter, Lewis Cope, says,

I have my own "rule of two." If someone makes some numerical claim, I look at the numbers, then see how much I might change the finding by adding or subtracting two from any of the figures. For example, someone says there are five cases of cancer in a community. Would it seem meaningful if there were three?

Or if there were eight cases this year but four the year before—a 100 percent increase—I ask myself, "If I add two cases to last year's total and subtract two from this year's, is there a chance things haven't changed, except by chance?" This approach will never supplant refined analysis. But by playing around with the numbers this way—I sometimes try three instead of two—a reporter can often spot a potential problem or error.

A statistician says, "This can help with small numbers but not large ones." Mosteller contributes "a little trick I use a lot on counts of any size." He explains, "Let's say some political unit has 10,000 crimes or deaths or accidents this year. Has something new happened? The minimum standard deviation [see

2503012490

page 33] for a number like that is 100—that is, the square root of the original number. That means the number may vary by a minimum of 200 every year without even considering growth, the business cycle, or any other effect. This will supplement your reporter's approach."

Looking for error in reported results, statisticians try to spot both false positives and false negatives. The *false positive* (or *Type I* or *alpha error* in statistical language you may see) is to find a result or effect where there is none. The *false negative* (or *Type II* or *beta error*) is to miss an effect where there is one. The latter is particularly common when there are small numbers. "There are some very well conducted studies with small numbers, even five patients, in which the results are so clear-cut that you don't have to worry about power," says Dr. Relman. "You still have to worry about applicability to a larger population, but you don't have to doubt that there was an effect. When results are negative, however, you have to ask, How large would the effect have to be to be discovered?"

Many scientific and medical studies are underpowered—that is, they include too few cases. "Whenever you see a negative result," another scientist says, "you should ask, What is the power? What was the chance of finding the result if there was one?" One study found that an astonishing 70 percent of 71 well-regarded clinical trials that reported no effect had too few patients to show a 25 percent difference in outcome. Half of the trials could not have detected a 50 percent difference.*

A statistician scanned an article on colon cancer in a leading journal. "If you read the article carefully," he said, "you will see that if one treatment was better than the other—if it would increase median survival by 50 percent, from five to seven and a half years, say—they had only a 60 percent chance of finding it out. That's little better than tossing a coin!"

The weak power of that study would be expressed numerically as .6, or 60 percent. Scan an article's fine print or footnotes, and you will sometimes find such a *power statement*. Most

authors still do
cially when re

How larg
lated that a tri
percent chanc

Sometimes
kind of cancer
pect that the :
X, you woul
excess rate to
significance. I
suffer a myoc
oral contrace
cent sure of o
you would ha

Even the
zero numerat
treated 14 leu
lar dysfunctio
remains. how
any re
may be

All this :
What's the size
20 individual
persons wou
Always try to

The mos
them. When
numbers anc
people, or ev

*And know
people Statistical
one or more per
scores, or physio
term science for

2503012491

2503012492

THE SCIENTIFIC WAY

authors still don't report one, but the practice is growing, especially when results are negative.

How large is a large enough sample? One statistician calculated that a trial has to have 50 patients before there is even a 30 percent chance of finding a 50 percent difference in results. Sometimes large populations indeed are needed.¹⁰ If some kind of cancer usually strikes 3 people per 2,000, and you suspect that the rate is quadrupled in people exposed to substance X, you would have to study 4,000 people for the observed excess rate to have a 95 percent chance of reaching statistical significance. The likelihood that a 30-to-39-year-old woman will suffer a myocardial infarction, or heart attack, while taking an oral contraceptive is about 1 in 18,000 per year. To be 95 percent sure of observing at least one such event in a one-year trial, you would have to observe nearly 54,000 women.¹¹

Even the lack of an effect — statistically sometimes called a zero numerator — can be a trap. Say, someone reports, "We have treated 14 leukemic boys for five years with no resulting testicular dysfunction" — that is, zero abnormalities in 14. The question remains, how many cases would they have had to treat to have any real chance of seeing an effect? The probability of an effect may be small yet highly important to know about.

All this means you must often ask, *What's your denominator?* What's the size of your population?¹² A disease rate of 10 percent in 20 individuals may not mean much. A 10 percent rate in 200 persons would be more impressive. A rate is only a figure. Always try to get both the numerator and the denominator.

The most important rule of all about any numbers: Ask for them. When anyone makes an assertion that should include numbers and fails to give them, when anyone says that most people, or even X percent, do such and such, you should ask,

*And know that to a statistician a population does not necessarily mean a group of people. Statistically, a population is any group or collection of pertinent units — units with one or more pertinent characteristics in common — people, events, objects, records, test scores, or physiological values (like blood pressure readings). Statisticians also use the term universe for a whole group of people or units under study.

What are your numbers? After all, some researchers reportedly announced a new treatment for a disease of chickens by saying, "33.3 percent were cured, 33.3 percent died, and the other one got away."

Bias and Confounders

One scientist once said that lefties are overrepresented among baseball's heavy hitters. He saw this as "a possible result of their hemispheric lateralization, the relative roles of the two sides of the brain." A critic who had seen more ball games said some simpler covariates could explain the difference. When they swing, left-handed hitters are already on the move toward first base. And most pitchers are right-handers who throw most often to right-handed hitters.¹²

Scientist A was apparently guilty of *bias*, meaning the introduction of spurious associations and error by failing to consider other influential factors. The other factors may be called *covariates*, *covariates*, *intervening or contributing variables*, *confounding variables*, or *confounders*. A simpler term may be "other explanations."

Statisticians call bias "the most serious and pervasive problem in the interpretation of data from clinical trials" . . . "the central issue of epidemiological research" . . . "the most common cause of unreliable data." Able and conscientious scientists try to eliminate biases or account for them in some way. But not everybody who makes a scientific, medical, or environmental claim is that skilled. Or that honest. Or that all-powerful. Some biases are unavoidable by the very difficulty of much research, and the most insidious biases of all, says one statistician, are "those we don't know exist."

Some biases may be uncovered by assiduous investigation. A father noticed that every time one of his 11 kids dropped a piece of bread on the floor, it landed with the buttered side up. "This utterly defies the laws of chance," he exclaimed. Close examination disclosed the cause: The kids were buttering their bread on both sides.

I told thi
called about .
prizes in a ch
that this coulc
bought nearly

He had c
tist and repor
factors?

Not ever
human failing
"I wouldn't
investigators
may be so e
over-rosy hu

Other px
motion and p
scious or unc
bias. Dr. Th
New York tr
firm,
main .
though not s
drugs for di
previously p
acknowledge
known to th

In cont
drug firm b
signed by ir
side board k
outcome. "It
interest in bi
disclosed so

Even a
Johns Hopk
with prisms

-2503012493

I told this story to one statistician, who said, "I was once called about a person who had won first, second, and third prizes in a church lottery. I was asked to assess the probability that this could have happened. I found out that the winner had bought nearly all the tickets."

He had of course asked the obvious question for both scientist and reporters: *Could the relationship described be explained by other factors?*

Not everyone will tell you, of course, for bias is a pervasive human failing. As one candid scientist is said to have admitted, "I wouldn't have seen it if I hadn't believed it." Enthusiastic investigators often tell us their findings are exciting. But they may be so exciting that the investigators paint the results in over-rosy hues.

Other powerful human drives—the race for academic promotion and prestige, financial connections—can also create conscious or unconscious conflicts of interest or attitudes that feed bias. Dr. Thomas Chalmers of Mount Sinai Medical Center in New York tells of a drug trial, financed by a pharmaceutical firm, in which both the head of the study committee and the main statisticians and analysts were the firm's employees, though not so identified in any credits. He tells of a study of oral drugs for diabetes in which the fact that the first author had previously published 14 articles on the subject, and in 7 had acknowledged support by the drug manufacturers, was "not known to the reader."

In contrast, Chalmers describes a study also financed by a drug firm but with a contract specifying a study protocol designed by independent investigators and monitored by an outside board less likely to be influenced by a desire for a favorable outcome. "It is never possible to eliminate" potential conflicts of interest in biomedical research, he concludes, but they should be disclosed so others can evaluate them.¹³

Even a genius may be biased. Horace Freeland Judson of Johns Hopkins University tells how Isaac Newton experimented with prisms and lenses and developed a theory of color, light,

2503012494

and the solar spectrum. He did not report seeing some dark lines—absorption lines, which mark varying wavelengths—that his instruments must have shown. A modern scientist argues that Newton's theory, not his instruments, had no place for that evidence: "To the observing scientist, hypothesis is both friend and enemy."¹⁴

For years technicians making blood counts were guided by textbooks that told them two or more "properly" studied samples from the same blood should not vary beyond narrow "allowable" limits. Reported counts always stayed inside those limits. A Mayo Clinic statistician rechecked and found that at least two thirds of the time the discrepancies exceeded the supposed limits. The technicians had been seeing what they had been told to expect and discounting any differences as mistakes. This also saved them from the additional labor of doing still more counting.

Both the *biased observer* and the *biased subject* are common in medicine. A researcher who wants to see a treatment result may see one. A patient may report one out of eagerness to please the researcher. There is also the powerful *placebo effect*. Summarizing many studies, one scientist found that half the patients with headaches or seasickness—and a third of those suffering from coughs, mood changes, anxiety, the common cold, and even the disabling chest pains of angina pectoris—reported relief from a "nothing pill."¹⁵ A placebo is not truly a nothing pill; the mere expectation of relief seems to trigger important effects within the body. But in a careful study the placebo should not do as well as a test medication; otherwise the test medication is no better than a placebo.

Sampling bias is the bugaboo of both political polls and medical studies. Say you want to know what proportion of the populace has heart disease, so you stand on a corner and ask people as they pass. Your sample is biased, if only because it leaves out those too disabled to get around. Your problem, a statistician would say, is *selection*. A political pollster who fails to build a valid probability sample, easy when questioning only a thousand or

so people fro
A doctor:
patient popul
average—ma
tion as a who
treat relativ
the dispropo
cally seek ou
Cleveland or
ber of difficu
affluent and
were valuabl
the samples
men and wo

An inve
distorting, a
otherwise th
in those dis
omits those
people
are dr
they came d
away, they c
had unfavor

Mostell
ous anesthet
hospitals. U
dead had b
plained by :
wound up v

The pr
ected, whe
of patients t
treated *cont*
compared. I
randomized

2503012495

so people from coast to coast, has equally poor selection.¹⁶

A doctor in a clinic or hospital with an unrepresentative patient population — healthier or sicker or richer or poorer than average — may report results that do not represent the population as a whole. Veterans Administration hospitals, for example, treat relatively few women; their conclusions may apply only to the disproportionate number of lower-income men who typically seek out the VA hospitals' free care. A celebrated Mayo or Cleveland or Ochsner clinic sees both a disproportionate number of difficult cases and a disproportionate number of patients affluent and well enough to travel. The famed Kinsey reports were valuable revelations of sexual behavior but flawed because the samples consisted disproportionately of upper middle-class men and women and of those willing to talk.

An investigator may also introduce bias by constraining, or distorting, a sample — by failing to reveal nonresponse or by otherwise "throwing away data." A surgeon cites his success rate in those discharged from the hospital after an operation but omits those who died during or just after the procedure. Many people drop out of studies — sometimes they just quit — or they are dropped for various reasons: They could not be evaluated, they came down with some "irrelevant" disorders, they moved away, they died. In fact, many of those not counted may have had unfavorable outcomes had they stayed in the study. Mosteller tells of a nationwide study of a possibly dangerous anesthetic. The investigators relied on autopsy results at 38 hospitals. Unfortunately, only about 60 percent of the relevant dead had been autopsied, and "anything could have been explained by the missing 40 percent, so that part of the study wound up with a handful of nothing."

The presence of significant nonresponse can often be detected, when reading medical papers, by counting the number of patients treated versus the number of untreated or differently treated *controls* — patients with whom the treated patients are compared. If the number of controls is strikingly greater in a randomized clinical trial (though not necessarily in an epidemic-

2503012496

logical or environmental study), there were probably many dropouts. A well-conducted study should describe and account for them. A study that does not may report a favorable treatment result by ignoring the fate of the dropouts—a confounding variable.

Age, gender, occupation, nationality, race, income, socioeconomic status, health status, and powerful behaviors like smoking are all possible confounding—and frequently ignored—variables. In the 1970s, foes of adding fluoride to city water pointed to crude cancer mortality rates in two groups of 10 U.S. cities. One group had added fluoride to water, the other had not, and from 1950 to 1970 the cancer mortality rate rose faster in the fluoridated cities. The National Cancer Institute pointed out that the two groups were not equal: The difference in cancer deaths was almost entirely explained by differences in age, race, and sex. The age-, race-, and sex-adjusted difference actually showed a small, unexplained lower mortality rate in the fluoridated cities.¹⁷

If you look carefully at the fate of women taking birth control pills, you find that advancing age and smoking are the two great confounders. You must take both into account to find the greatest clusters of ill effects. Smoking has been an important confounder in studies of industrial contaminants like asbestos, in which, again, the smokers suffer a disproportionate number of ill effects.¹⁸

A 1947 survey of Chicago lawyers showed that those who had mere high school diplomas before entering legal training earned 6.3 percent more, on the average, than college graduates. The confounder here—the real explanation—was age. In 1947 there were still many older lawyers without college degrees, and they were simply older, on the average, and hence more established.¹⁹

Occupational studies often confront another seeming paradox: The workers exposed to some possible adverse effect turn out to be healthier than a control group of persons without such exposure. The confounder: the well-known *healthy-worker effect*.

Workers tend
in general.

Some stu
increase in car
gens. It took a
They commo
were emitted.
stratified, or br
Such findings
genetics, whe
ing or ruling
bles—are occ
put blacks in
ent variable, th

In a tw
plains, "one
which affects
that more pe
seen as the i
incident
of course,
stantly expo
than others.
the black w
one indepen
portant und
may be that
each other,
coworkers, t
cold weather
drying nasa
viruses.

The se
pursuits of
physician w
any student

2503012497

Workers tend to be healthier and live longer than the population in general.

Some studies of workers in steel mills showed no overall increase in cancer, despite possible exposures to various carcinogens. It took a look at black workers alone to find excess cancer. They commonly worked at the coke ovens, where carcinogens were emitted. This was a case where the population had to be *stratified*, or broken up in some meaningful way, to find the facts. Such findings in blacks often may be falsely ascribed to race or genetics, when the real or at least the most important contributing or ruling variables—to a statistician, the *independent variables*—are occupation and the social and economic plights that put blacks in vulnerable settings. The excess cancer is the *dependent variable*, the result.

"In a two-variable relationship," Dr. Gary Friedman explains, "one is usually considered the independent variable, which affects the other or dependent variable."²⁰ Take the fact that more people get colds in winter. Here weather is commonly seen as the underlying or independent variable, which affects incidence of the common cold, the dependent variable. Actually, of course, some people, like children in school who are constantly exposed to new viruses, are more vulnerable to colds than others. In the case of these children, then, as in the case of the black workers at the coke ovens, there is often more than one independent variable. Also, some people think that an important underlying reason for the prevalence of colds in winter may be that children are congregated in school, giving colds to each other, thence to their families, thence to their families' coworkers, thence to the coworkers' families, and so on. But cold weather—and home heating?—may still figure, perhaps by drying nasal passages and making them more vulnerable to viruses.

The search for *true variables* is obviously one of the main pursuits of the epidemiologist, or disease detective—or of any physician who wants to know what has affected a patient, or of any student of society who seeks true causes. Like colds, many

2503012498

medical conditions, such as heart disease, cancer, and probably mental illness, have multiple contributing factors. Where many known, measurable factors are involved, statisticians can use mathematical techniques—the terms you will see include *multiple regression*, *multivariate analysis*, and *discriminant analysis and factor, cluster, path, and two-stage least-squares analyses*—to relate all the variables and try to find which are the truly important predictors. Yet some situations, like the striking decline in U.S. heart disease mortality in recent years, defy such analyses. These years have seen several major changes in American life that may play a role: less smoking among men, consumption of a leaner diet, more recreational exercise (though more sedentary work). Medical care is far better, including the treatment of hypertension, which disposes people to heart disease. Many of these variables cannot be well measured, and the effect of some is debatable, so—a common situation in science—the truth remains uncertain.

Variability

Doctors always say, "Most things are better in the morning," and they're mostly right. Most chronic or recurring conditions wax and wane. We tend to wake up at night when the condition is at its worst. Then, no matter what is done by way of treatment the next day, the odds are that we'll feel better.

This is *regression toward the mean*: the tendency of all values in every field of science—physical, biological, social, and economic—to move toward the average. Tall fathers tend to have shorter sons, and short fathers, taller sons. The students who get the highest grades on an exam tend to get somewhat lower ones the next time. The regression effect is common to all repeated measurements.

Regression is part of an even more basic phenomenon: *variation*, or *variability*. Virtually everything that is measured varies from measurement to measurement. When repeated, every experiment has at least slightly different results. Take a patient's

blood pressure now, and the reading will be different at different times; it may vary greatly.

The importance of variability also measures and observer doctors will not be grossly different. heart murmur hearing to determine one time to the cancer researcher usual regular too well and not enough variation.

Biological physiology patients, react differ in variations, within the same

Every patient each with its own such as height and—if we take action—we may. We can't get need single

Enter the median, and some idea of its properties, etc.

When we take the mean or arithmetic number of

2503012499

blood pressure, pulse rate, or blood count several times in a row, and the readings will be somewhat different. Take them at different times of day or on different days; and the readings may vary greatly.

The important reasons? In part, fluctuating physiology, but also measurement errors, the limits of measurement accuracy, and observer variation. Examining the same patient, no two doctors will report exactly the same results, and the results may be grossly different. If six doctors examine a patient with a faint heart murmur, only one or two may have the skill or keen hearing to detect it. Experimental results so typically differ from one time to the next that scientific and medical fakers—a Boston cancer researcher, for example—have been detected by the unusual regularity of their reported results, with numbers agreeing too well and the same results appearing time after time, with not enough variation from patient to patient.

Biological variation is the most important cause of variation in physiology and medicine. Different patients, and the same patients, react differently to the same treatment. Disease rates differ in different parts of the country and among different populations, and—alas, nothing is simple—there is natural variation within the same population.

Every population, after all, is a collection of individuals, each with many characteristics. Each characteristic, or *variable*, such as height, has a *distribution* of values from person to person, and—if we would know something about the whole population—we must have some handy summaries of the distribution. We can't get much out of a list of 10,000 measurements, so we need single values that summarize many measurements.

Enter here the familiar *average* or, more exactly, the *mean*, *median*, and *mode*. These and a few other measures can give us some idea of the look of the whole and its many measurable properties, or *parameters*.

When most of us speak of an average, we mean simply the *mean* or *arithmetic average*, the sum of all the values divided by the number of values. The mean is no mean tool; it is a good way

2503012500

to get a typical number, but it has limitations, especially when there are some extreme values. There is said to be a memorial in a Siberian town to a fictitious Count Smerdlovski, the world's champion at Russian roulette. On the average he won, but his actual record was 73 and 1.²¹

If you look at the average salary in a hospital, you will not know that half the personnel may be working for the minimum wage, while a few hundred persons make \$100,000 or more a year. You may learn more here from the *median*, the figure that divides a population into two equal halves. The median can be of value when a group has a few members with extreme values, like the 400-pounder at an obesity clinic whose other patients weigh from 180 to 200 pounds. If he leaves, the patients' mean weight might drop by 10 pounds, but the median might drop just 1 pound.²²

The most frequently occurring number or value in a distribution is called the *mode*. When the median and the mode are about the same, or even more when mean, median, and mode are roughly equal, you can feel comfortable about knowing the typical value.

You still need to know something about the exceptions, in short, the *dispersion* (or spread or scatter) of the entire distribution. One measure of spread is the *range*. It tells you the lowest and highest values. It might inform you, for example, that the salaries in that hospital range from \$10,000 to \$250,000.

You can also divide your values into 100 *percentiles*, so you can say someone or something falls into the 10th or 71st percentile, or into *quartiles* (fourths) or *quintiles* (fifths). One useful measure is the *interquartile range*, the interval between the 75th and 25th percentiles—this is the distribution in the middle, which avoids the extreme values at each end. Or you can divide a distribution into *subgroups*—those with incomes from \$10,000 to \$20,000, for example, or ages 20 to 29, 30 to 39, and so on.

All these values can easily be plotted. With many of the things that scientists, economists, or others measure—IQs, for example, and other test scores—we typically tend to see a famil-

iar, bell-shaped end, or *tail*. T 19th-century (But you may clusters, a *bim*

A widely great deal. Nc tance from th range, this ha how spread o In what one s in most sets · being measur average by n more than 2 than 2.57 sta

"Once yo shaped distri the whole ric curve variation of t the more spr

*There is n depending on the differences between number of square of a population re result. As in

Sometimes mean, being an standard error or a

2503012501

All the abo

iar, bell-shaped *normal distribution*, high in the middle, low at each end, or *tail*. This is the classic *Gaussian curve*, named after the 19th-century German mathematician Karl Friedrich Gauss. But you may also find that the plot has two or more peaks or clusters, a *bimodal* or *multimodal distribution*.

A widely used number, the *standard deviation*, can reveal a great deal. No matter how it sounds, it is not the average distance from the mean but a more complex figure.* Unlike the range, this handy figure takes full account of every value to tell how spread out things are—how dispersed the measurements. In what one statistician calls a truly remarkable generalization, in most sets of measurement “and without regard to what is being measured” only 1 measurement in 3 will deviate from the average by more than 1 standard deviation, only 1 in 20 by more than 2 standard deviations, and only 1 in 100 by more than 2.57 standard deviations.

“Once you know the standard deviation in a normal, bell-shaped distribution,” according to Thomas Louis, “you can draw the whole picture of the data. You can visualize the shape of the curve without even drawing the picture, since the larger the variation of the numbers, the larger the standard deviation and the more spread out the curve—and vice versa.”

*There is more than one way to calculate it, and there are several variations, depending on the statistician's purpose. A common one is to add the squares of the differences between each number and the mean, then divide that number by the total number of squares, often referred to as the *variance* (minus 1 if you're looking at a sample of a population rather than the whole population). Then calculate the square root of the result. As in

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

Sometimes statisticians calculate the *standard deviation of the mean*—this because the mean, being an average, is less variable than single measurements. Some call this the *standard error of the mean*. As in

$$\sigma_x = \frac{1}{\sqrt{n}}$$

All the above are measures of dispersion.

2503012502

Example: If the average score of all students who take the SAT college entrance test is relatively low and the spread—the standard deviation—relatively large, this creates a very long-tailed, low-humped curve of test scores, ranging, say, from around 300 to 1500. But if the average score of a group of brighter students entering an elite college is high, the standard deviation of the scores will be less and the curve will be high-humped and short-tailed, going from maybe 900 to 1500.

"If I just told you the means of two such distributions, you might say they were the same," another scientist says. "But if I reported the means and the standard deviations, you'd know they were different, with a lot more variations in one."

From a human standpoint, variation tells us that it takes more than averages to describe individuals. Biologist Stephen Jay Gould learned in 1982 that he had a serious form of cancer. The literature told him the median survival was only eight months after discovery. Three years later he wrote in *Discover*, "All evolutionary biologists know that means and medians are the abstractions," while variation is "the reality," meaning "half the people will live longer" than eight months.

Since he was young, since his disease had been diagnosed early, and since he would receive the best possible treatment, he decided he had a good chance of being at the far end of the curve. He calculated that the curve must be skewed well to the right, as the left half of the distribution had to be "scrunched up between zero and eight months, but the upper right half [could] extend out for years." He concluded, "I saw no reason why I shouldn't be in that small tail. . . . I would have time to think, to plan and to fight." Also, since he was being placed on an experimental new treatment, he might if fortune smiled "be in the first cohort of a new distribution with . . . a right tail extending to death by natural causes at advanced old age."²³

Statistics cannot tell us whether fortune will smile, only that such reasoning is sound.

Studie. Good

Why think? Why

Sit down before &
notion, follow hui
shall learn nothin

This is the part I

T HERE is
cally than Alz
the brain. At
team imagina
tients aged 64
theoretically p
solution for c

After 18
that accordin
patients show
held his own.
the work and
do a story ar

2503012503

Studies, Good and Bad

4

Why think? Why not try an experiment?

—John Hunter
1728-1793 British anatomist

Sit down before fact as a little child, be prepared to give up every preconceived notion, follow humbly wherever and to whatever abysses nature leads, or you shall learn nothing.

—Thomas Henry Huxley

This is the part I always hate.

—A mathematician as he approaches the equal sign (in a Sidney Harris cartoon in *American Scientist*)

THERE is no disease that strikes older people more tragically than Alzheimer's disease, which makes a useless tangle of the brain. At a prestigious New England university a research team imaginatively inserted catheters into the skulls of four patients aged 64 to 73 to deliver a continuous infusion of either a theoretically promising drug or, alternately, an ineffectual saline solution for comparison.

After 18 months the investigators published a paper saying that according to observations by the patients' families, three patients showed marked improvement and the fourth at least held his own. Fascinating, of course. Some reporters learned of the work and began inquiring. The investigators let a TV crew do a story and also held a news conference, with one patient

35

2503012504

brought forth for on-camera testimonials. Except for some newspapers that decided to print nothing, the story flew far and wide.

The head investigator, a chief resident in neurosurgery, cautioned that the results, though encouraging, were "very early" and "certainly do not prove this is an effective treatment." He advised healthy skepticism. But headlines unequivocally read: "Alzheimer's Test Found Successful," "Alzheimer's: A New Promise," "First Breakthrough Against Alzheimer's," "Pump Offers Hope," "Possible Alzheimer's Cure."

Within two months the medical center logged 2,600 phone calls, mainly from desperate families, and critics began asking why a press conference had been held, since a study of only four patients—with unblinded investigators getting their assessments from hopeful families—meant little.

Harvard's Dr. Jay Winsten concluded that "the decision to hold a press conference . . . far outweighed in impact the modulating effect of the investigators' qualifying language. The visual impact of [one] patient's on-camera testimonials all but guaranteed that TV coverage would oversell the research, despite any qualifying language."¹

When dubious claims are made—about Alzheimer's, a new cancer drug, a possible AIDS cure—and the claims get widely reported, there is commonly a lot of postmortem clucking and soul-searching among reporters and editors. Then someone else makes some sensational claim, and the same thing may happen all over again.

The biggest error in medical science, according to Dr. Thomas Chalmers, is "the uncontrolled pilot study in which the investigators try a treatment on 10 patients, and if it seems to work . . . are tempted to report it" to fellow scientists, let alone the media.²

All science is only a stab at the truth. Even with the best of statistics, "We scientists don't know how to tell the whole truth," Mosteller reminds us.³ Outside this honest limitation lie vast realms of inadequate science with plausible-sounding yet shaky

statistics. A Frenchman said 150 years ago that numerical in time than the real often give it. "Science every idiot in the program thinks

The big problem have little to do with judgment to conduct it, the frenzied media many chances calls for sophis hope of telling report?

A funda ducted studi include and to th methods, and this kind of ex

This is r. there is much of numbers :

Experimentation

Student credit-rated, what has been studies car

Science as generaliz tured into science. Of

N 03012505

statistics. A French physician, Pierre Charles Alexandre Louis, said 150 years ago, "The only reproach which can be made to the numerical method" is that it "requires much more labor and time than the most distinguished members of our profession" often give it. "Some days," says one modern statistician, "I think every idiot in the country who can put his hands on a computer program thinks he's a statistician."

The big problems of statistics, say its best practitioners, have little to do with computations and formulas. They have to do with judgment, we're told, with how to design a study, how to conduct it, then analyze and interpret the results. In a day of frenzied media competition for the public's eye and ear—and many chances to do harm by shaky reporting—journalism too calls for sophisticated judgment. How, then, can we have some hope of telling which studies seem credible, which we should report?

A fundamental principle is that every conscientiously conducted study has a careful *design*: a method or plan of attack to include the right kind and number of patients or petri dishes and to try to eliminate bias. Different problems require different methods, and one of the most basic questions in science is, *Can this kind of experiment, this design, yield the answer?*

This is not a simple question for a reporter to answer, but there is much we can know. What kinds of studies, what kinds of numbers and controls and methods, should we look for?

Experiments versus Seductive Anecdotes

Students and eggs can be graded, citizens and cities can be credit-rated, and scientific evidence can be weighed according to what has been called a hierarchy of evidence. Some kinds of studies carry little weight, some more, some a great deal.

Science and medicine started with *anecdotes*, unreliable as far as generalization is concerned, yet provocative. Anecdotes matured into systematic *observation*, the most ancient form of science. Observation told the ancients much about the stars, it

2503012506

told the pharaohs' physicians much about the sick, and it is still important, for simple "eyeballing" has developed into *data collection* and the recording of *case histories*. These are respectable, yea, indispensable methods yet still only one part of science. Case histories may not be typical, or they may reflect the beholder. Medicine continues to be plagued by Big Authorities who insist, "I know what I see."

There can be useful, even inspired, observation and analysis of *natural experiments*. Excess fluoride in some waters hardened teeth, and this observation led to fluoridation of drinking water to prevent tooth decay. There are also man's inadvertent experiments, disastrous and benign, to be studied. Hiroshima triggered wide analysis of the effects of nuclear radiation, invaluable yet frustrating because there were no good measures of exposure levels, a gap that has caused confusion and controversy ever since.

In 1585 or so, Galileo dropped those weights from a tower and helped invent the *scientific experiment*: a study in which the experimenter *controls* the conditions—controlled conditions are the heart of the experimental method—and records the effect. Experiments on objects, animals, germs, and people matured into the modern *experimental study*, in which the experimenter typically changes only one or some other planned number of variables to see the outcome.

Clinical Trials

The experimental method is the essence of experimental medicine's current "gold standard":⁴ the *controlled, randomized clinical trial*. At its best, the investigator tests a treatment or drug or some other intervention by randomly selecting at least two comparable groups, the *experimental group* that is tested or treated and a *control group* that is observed for comparison.

True clinical trials are expensive and difficult. It has been estimated that of 100 scheduled trials, 60 are abandoned, not

implemented, occult in recruit
ments, or, some
(making contr
group unethica
sults, and just
theless are calle
to evaluate in
Randomized c
heart attack de
strokes, and t
No doctor, ob
shown these ti

Types of
• Among
similar group
no treatment.

• In cross
ments in succ
controlle
observa
treatment. T
outcome of ti
between stuc
become mor
health-consci
patients in :
studies eithe
cholesterol a
some of the
fewer fats—

• Inves
son with o
percent, sa
eternal contr

2503012507

2503012508

STUDIES, GOOD AND BAD

39

implemented, or not completed, whether for lack of funds, difficulty in recruiting or keeping patients, toxicity or other problems, or, sometimes, rapid evidence of a difference in effect (making continued denial of effective treatment to a control group unethical). Another 20 trials produce no noteworthy results, and just 20, results worth publishing. Clinical trials nonetheless are called the strongest, most precise, most decisive way to evaluate medical interventions and learn true causation. Randomized clinical trials proved that new drugs could cut the heart attack death rate, that treating hypertension could prevent strokes, and that polio, measles, and hepatitis vaccines worked. No doctor, observing a limited number of patients, could have shown these things.

Types of clinical studies include the following:

- Among the most reliable are *parallel studies* comparing similar groups given different treatments, or a treatment versus no treatment. But such studies are not always possible.
- In *crossover studies* the same patients get two or more treatments in succession and act as their own controls. Similarly, *self-controlled studies* evaluate an experimental treatment by control observations during periods of no treatment or of some standard treatment. There are pitfalls here. Treatment A might affect the outcome of treatment B, despite the usual use of a *washout period* between study periods. Patients become acclimated: They may become more tolerant of pain or side effects or, now more health-conscious, may change their ways. The controls—the patients in a control group—don't always behave in parallel studies either: In one large-scale trial of methods to lower blood cholesterol and risk of heart disease, many controls adopted some of the same methods—quitting cigarette smoking, eating fewer fats—and reduced their risk too.
- Investigators often use *historical controls* (meaning comparison with old records: historically the cure rate has been 30 percent, say, and the new therapy cures 60 percent) or other *external controls* (such as comparison with other studies).

controls are often misleading—the groups compared are frequently not comparable, the treatments may have been given by different methods—but they are still at times useful.

What Makes a Study Honest?

Obviously, all studies, including the best, have potential pitfalls:

- *Lack of adequate controls* is fatal if you really want to put the results in the bank.
- The group or sample studied, 10 people or 10,000, must be large enough to get a valid result and representative enough to apply to a larger population. Because people vary so widely in their reactions, and a few patients can fool you, fair-sized groups of patients are usually needed. And enough of the right kind of subjects are needed for a suitable sample. Picking patients for a medical study is no different from picking citizens to be questioned in a political poll. In both, a sample is studied, and *inferences*—the outcome of an election, the results in patients in general—are made for a larger population.

To get a large enough sample, medical researchers more and more try to conduct *multicenter trials*, which are appealing because they can include hundreds of patients, but expensive and tricky because one must try to maintain similar patient selection and quality control at 10 or 100 institutions. Successful multicenter trials established the value of controlling hypertension to prevent strokes. They demonstrated the strong probability that less extensive surgery is as effective as more drastic surgery for many breast cancers.

- The sample should be randomized—divided by some random method into comparable experimental and control groups. Randomization can easily be violated. A doctor assigning patients to treatment A or B may, seeing a particular type of patient, say or think, "This patient will be better on B."

If treatment B has been established as better than A, there should be no random study in the first place and certainly no

study of that c
"the trial's gu
one critique. A
are often assi
puter-generat

• To comb
and get answe
study popula
groups by ag
stratify can h
tampons in t
cases were bi

The ide
can be trick
may fail to s
But some p
stronger pat
treated imm

We rep
or dir
nomic
major news;
parity with
vantaged gr
page, did t
older peopl
incomes be
are still ma

• To co
blinded—to
blinded, so i
a treatmen
know whe
ineffective;
a good res
There is a

2503012509

study of that doctor's patient. When randomization is violated, "the trial's guarantee of lack of bias goes down the drain," says one critique. As a result, patients who consent to randomization are often assigned to study groups according to a list of computer-generated random numbers.

• *To combat bias*—the influence of confounding variables—and get answers applicable to various populations, the sample or study population must often be *stratified*, or separated into groups by age, sex, socioeconomic status, and so on. Failure to stratify can hide true associations. The role of high-absorbency tampons in toxic shock syndrome was clarified only when the cases were broken down by precise type of tampon used.

The identification of important subcategories of patients can be tricky indeed. A study of open-heart surgery patients may fail to separate out those who had to wait for their surgery. But some patients die waiting, and those left are relatively stronger patients who do better, on the average, than those treated immediately after diagnosis.

We reporters may also fail to pay attention to stratification, or distribution. In early 1985 the President's Council of Economic Advisers reported that—to quote the page-one lead in a major newspaper—"elderly Americans have achieved economic parity with the rest of the population and no longer are a disadvantaged group." Not for several paragraphs, now on an inside page, did the story note that "there's a lot of variability," and older people are also "more likely . . . to have members with incomes below the average of their age group." In short, there are still many elderly trapped in poverty.

• *To combat bias in investigators or patients*, studies should be *blinded*—to the extent feasible, *single*-, *double*-, or, best of all, *triple-blinded*, so that neither the doctors nor the nurses administering a treatment nor the patients nor those who assess the results know whether today's pill is treatment A, treatment B, or an ineffective *placebo*. Otherwise, a doctor or patient who yearns for a good result may see or feel one when the "right" drug is given. There is a tale of an overzealous receptionist who, knowing

2503012510

which patients were getting the real drug and not the placebo, was so encouraging to these patients that they began saying they felt good, willy-nilly.*

Barring observant receptionists, the use of a placebo—from the Latin meaning “I shall please”—may help maintain blindness. Placebos actually give some relief in a third of all patients, on the average, in various conditions. The effect is usually temporary, however, and a truly effective drug ought to work substantially better than the placebo.

Blinding is often impossible or unwise. Some treatments don't lend themselves to it, and some drugs quickly reveal themselves by various effects. But an unblinded test is a weaker test.

- Finally, what makes a study honest is honesty. John Bailar warns of deliberate or careless deceptions that seem to be universally accepted today, practices that sometimes have much value but at other times are “inappropriate and improper and, to the extent that they are deceptive, unethical.” Among them: the selective reporting of findings, leaving out some that might not fit the conclusion; the reporting of a single study in multiple fragments, when the whole might not sound so good; and the failure to report the low power of some studies, their inability to detect a result even if one existed.⁷

Dr. Charles Moertel of the Mayo Clinic says,

Probably the majority of cancer patients treated with chemotherapy today are receiving regimens that have not been proved effective by randomized trial. . . . Many articles published in our major journals make claims for fantastic therapeutic accomplishments with no randomized controls. . . . Many, if not most, of the randomized studies . . . are of such poor quality that their results are unbelievable. . . . Precious few have withstood the scrutiny of carefully designed confirmatory scientific study.

He calls a multitude of poor methods statistical legerdemain: “the games we play, trying to squeeze out that little bit of breakthrough.” Why the pressure to play them? “Salvation,” Dr.

David Salsbury
prestige, invit
references in 1

Epidemiology

Clinical :
populations,
epidemiology se
a population
cal investigatio

Epidemiologi
ies—some si
same pitfalls
the right ans
goes, an epi
sex.

Epidemiolog
epidemics o
molec
we li
the healthie
the first er
healthier to
today's envi
may have l
he might h
wealthier a

In 17
success by
don's chim
to soot—b
rette. A c
cases on :
drinking v
Street pu

2503012511

David Salsburg answers. "Fruit in this world (increases in salary, prestige, invitations to speak) and beyond this life (continual references in the citation index)."¹⁸

Epidemiology: Hippocrates to AIDS

Clinical studies deal with patients. Epidemiology deals with populations, which sometimes are large groups of patients. Epidemiology seeks the causes of both health and disease by placing a population under its own kind of microscope, the *epidemiological investigation*.

Epidemiological studies in many ways parallel clinical studies—some studies are both—and are subject to many of the same pitfalls and rules, like avoiding bias and stratifying to get the right answers about the right subgroups. An old saw, in fact, goes, an epidemiologist is a physician broken down by age and sex.

Epidemiology in its early days was concerned wholly with epidemics of typhoid, smallpox, and other infections. But epidemiologists today also ask, "What should we eat and how should we live to stay healthy?" and they study large groups to see how the healthiest and unhealthiest live. Hippocrates has been called the first environmentalist because he observed that it was healthier to live in high places than in low ones. Anticipating today's environmentalists, he blamed bad air and bad water and may have been partly right. But he failed to stratify; otherwise he might have noticed that the people who lived high were also wealthier and better nourished than those who lived low.¹⁹

In 1740 Percival Pott scored a famous epidemiological success by observing the high rate of scrotum cancer in London's chimney sweeps and correctly blaming it on their exposure to soot—burned organic material, much like a smoked cigarette. A century later, John Snow, plotting London cholera cases on a map and noting a cluster around one source of drinking water, removed the handle from the now famed Broad Street pump and helped end a deadly epidemic. The 19th-

2503012512

century French advocate of statistical methods, Pierre Louis, observed hospital patients and helped stop the use of bleeding as a treatment. Ignaz Semmelweis showed that doctors' dirty hands transmitted deadly childbed fever to mothers.

Modern epidemiologists successfully indicted smoking as a cause of lung cancer and heart disease and identified the association of fats and cholesterol with clogging of the arteries. They evaluate vaccines, assess new methods of health care delivery, and track down the causes of new scourges like AIDS, toxic shock syndrome, and Legionnaires' disease, all by several methods. All are valuable. All are full of traps.

- Epidemiology, like all of science, started with *observational* studies, and these remain important. They are weak and uncertain, we have noted, when it comes to determining cause and effect. Yet observation is how we first learned of the unfortunate effects of toxic rain, Agent Orange, cigarette smoking, and many sometimes helpful, sometimes harmful medications—and of certain sexual practices and addicts' use of dirty needles on AIDS.

- Some observational studies are simply *descriptive*—describing the incidence, prevalence, and mortality rates of various diseases, for example. Other, *analytic* studies seek to analyze or explain: the Seven-Country Study, for example, that helped associate high meat and dairy fat and cholesterol consumption with excess risk of coronary heart disease. *Ecological* studies look for links between environmental conditions and illness. Human migrations—like that of the Japanese who come to the United States, eat more fat, and get more disease than they did in Japan—are among valuable *natural experiments*.

- The simplest observational measurement is a *count*. *Sampling* is just a more sophisticated kind of count. You can't count or question everybody, so you seek a sample that represents the whole. Many epidemiological surveys rely on samples—among them, government surveys of health and nutritional habits. Samples and surveys often use *questionnaires* to get information.

A sample or survey is never more than a snapshot of the

scene at the moment unless frequent than the quality compared patients with those their almost half of time of a year. And people tend to often say both A survey may get accurate if

- Epidemiological control studies, cross-sectional or *case-control* studies look at the rates of effects by age in a study: A cross-section survey may last a few days.

A ^Y
a disease
examined groups
drome, mainly
ease Control
tients, or ease
their families
ries that cover
group is their
compeer group,
and other etiolog

The results
the case-control
tively easy, I
semble clues
may test for
use of tampons
as the main

2503012513

scene at the moment; it can't portray an ever-changing picture unless frequently repeated. Questionnaires may be no better than the quality of the answers, written or verbal. One survey compared patients' reporting of their current chronic illnesses with those their doctors recorded. The patients failed to mention almost half of the conditions the doctors detected over the course of a year. And whether it comes to illness, diets, or drinking, people tend to put themselves in the best possible light. They often say both yes and no to the same question in different form. A survey may stand or fall on the use of sophisticated ways to get accurate information.

- Epidemiologists' studies may also be *prevalence studies*, *case-control studies*, or *cohort studies*. A *prevalence study*, also called a *current control study*, or *cohort study*, is a wide-angle snapshot of a population: a *cross-sectional study* look at the rate of disease X or at toxic agent X and its possible effects by age, sex, or other variables. A political poll is such a study: A cross section of the nation is examined in a period of a few days.

A *case-control study* examines *cases* and *controls* for a close-up of a disease's relationship to other factors in a small, intensively examined group. The nation hears of cases of toxic shock syndrome, mainly in young women. The Federal Centers for Disease Control launches a *field investigation* to find a series of patients, or *cases*, confirm the diagnosis, then interview them and their families and other contacts to assemble careful case histories that cover, hopefully, all possible causes or associations. This group is then compared with a randomly selected but matched *control group*, or *control group*, of healthy young women of like age and other characteristics.

The results need to be interpreted with great caution, but the *case-control study* is often a quick, highly useful and relatively easy, low-cost first approach or fishing expedition to assemble clues about causes or even a working hypothesis. Or it may test some hypothesis. A *case-control study* pinpointed the use of tampons (later found to be certain high-absorbency ones) as the main villain in toxic shock. The relationship of cigarette

2503012514

smoking to lung cancer, the association of birth control pills with blood vessel problems, and the transmission patterns of AIDS were identified in case-control studies that pointed to the need for broader investigation.

Cohort or *incidence* studies are motion pictures. They pick a group of people, or *cohort*—a cohort was a unit of a Roman legion—often stratify or divide them into subgroups, then follow them over time, often for years, to see how some disease or diseases develop. These studies are costly and difficult. Subjects drop out or disappear. Large numbers must be studied to see rare events. But cohort studies can be powerful instruments and substitutes for randomized experiments that would be ethically impossible. You can't ethically expose a group to an agent that you suspect would cause a disease. You can watch a group so exposed.

The noted Framingham study of ways of life that might be associated with developing heart disease has followed more than 5,000 residents of that Massachusetts town since 1948. The American Cancer Society's 1952–55 study of 187,783 men aged 50 to 69, with 11,780 of them dying during that period, did much to establish that cigarette smoking was strongly associated with developing lung cancer.¹⁰

- Many epidemiological, as well as clinical, studies are handicapped because they must be *retrospective*. They look back in time—at medical records, vital statistics, or people's recollections (for example, those collected in interviews in a case-control study). People who have a disease are questioned to try to find common habits or exposures. Women with cervical cancer are interviewed to see how many took possibly guilty hormones and how many did not. People who live around a Love Canal are asked if they have been ill.

Retrospective studies are notoriously unreliable. Memories fail or play tricks. Old records are poor and misleading. Definitions of diseases and methods of diagnosis vary sharply over the years. The patients you find may not be representative. A retrospective study, however intriguing, generally only says that there may be something here that ought to be investigated.

(There are exce-
tive study can
lected in the pa
was a retrospec

- A prospec-
the American C
sharply on a s
statistical and :
ford tells how f
the accuracy c
adequate pros
ward looks we

- Epidem
experiments o
cally *interventio*
tion; somethir

The mas
Salk polio vac
trial too, with
to either
placeb
divided betw
first- and thi
participating
counted all c
those who h
In the place
the vaccinat
subjects late
shot.¹²

Anothe
tablished th
tooth decay
not. Bländir
tal caries th
cebo effect.

2503012515

STUDIES, GOOD AND BAD

47

(There are exceptions. Dr. Gary Friedman writes, "A retrospective study can be quite reliable if based on data carefully collected in the past. A revealing study of mortality in radiologists was a retrospective cohort study based on good data.")

- A prospective study, in contrast — like the Framingham and the American Cancer Society studies — looks forward. It focuses sharply on a selected group who are all followed by the same statistical and medical techniques. Dr. Eugene Robin at Stanford tells how four separate retrospective clinical studies affirmed the accuracy of a test for blood clots in the lungs. When an adequate prospective clinical trial was done, most of the backward looks were proved wrong.¹¹

- Epidemiology also includes *experimental studies*, the classical experiments of science on a larger human scale. These are typically *intervention studies*. There is some intervention or manipulation; something is done to some of the subjects.

The massive and hugely successful 1954 field trial of the Salk polio vaccine was a classic intervention trial and a clinical trial too, with 401,974 first- to third-graders assigned at random to either a vaccinated group or a control group injected with a placebo, or dummy shot — and another 947,171 children divided between vaccinated second-graders and unvaccinated first- and third-graders acting as controls. In addition, in all participating states or counties, the investigators studied and counted all cases of polio in a grand total of 1,829,916 children: those who had taken part in the study and those who had not. In the placebo areas, the study was also triple-blinded: neither the vaccinators, the subjects, nor the doctors who examined the subjects later for polio knew which children got which kind of shot.¹²

Another successful intervention study, a *community trial*, established the value of fluoridating water supplies to prevent tooth decay. Some towns had their water fluoridated; some did not. Blinding was impossible, but the striking difference in dental caries that resulted could not have been caused by any placebo effect.

Questions Reporters Can Ask

5

Just because Dr. Famous or Dr. Bigshot says this is what he found doesn't mean it is necessarily so.

-Dr. Arnold Reisman

Ask to see the numbers, not just the pretty colors.

- Dr. Richard Maryolin

National Institutes of Health,
describing PET scans to reporters.

WHAT questions should we reporters ask—to make our news solid, to report the more valid claims and ignore the weak and phony? When a scientist or physician or anyone else says, "I've discovered that . . ." what should we ask?

In 1949, a year after Britain's National Health Service—"socialized medicine"—was launched, my editors sent me to Britain to see how it was working. A bit stumped, I asked Dr. Morris Fishbein, the provocative genius who long edited the *Journal of the American Medical Association*, "How can I, a reporter, tell whether a doctor is doing a good job?" He immediately said, "Ask him how often he has a patient take off his shirt."

His lesson was plain: No physical examination is complete unless the patient takes off his or her clothes. Most reporters are not skilled statisticians, but we can ask some similarly revealing questions. Many of these are not even statistical, just simple ones that, like Fishbein's, probe soft spots and often disclose either a conscientious approach or one that can't be trusted.

We can learn here from one method of science. We said

18

earlier that a person seeking truth, often sees whether or much like the lawcutor to prove guilty. A reporter should be equal words or thought.

If an investor case, you may I since a good scientist for you. The something.

Here are some simple and obvious things you want to ask the

*How do you
men? WI*

*What kin.
design? And a.*

*What was
case-control, pre-
ter for kinds
people just s
conclusion w
medical edito
study? What
answering?*

250301251



earlier that a properly skeptical scientist, starting a study and seeking truth, often begins with a *null hypothesis*—that treatment A is no better than treatment B, that there's nothing there—then sees whether or not the evidence disproves it. This approach is much like the law's presumption of innocence: It is for the prosecutor to prove beyond reasonable doubt that the suspect is guilty. A reporter, without being cynical and believing nothing, should be equally skeptical and greet every claim by saying, in words or thought, "Show me."

If an investigator or claimant is competent and has a good case, you may have to ask none or very few of these questions, since a good scientific presentation should answer most of them for you. The need for a lot of questions could itself tell you something.

Here are some possible questions, then, some of them simple and obvious ones, a few more technical for those who might want to ask them.

How do you know? Have you done a study? Was there an experiment? What is the evidence? Or is the approach just anecdotal? Answers like "In my experience . . ." "In my hands . . ." "I've seen 20 cases . . ." and "There are four cases in our block . . ." may be interesting, may be worth scientific investigation, may be worth a cautious news story, but there is not yet anything like certainty.

What kind of study was it? Was there a systematic research plan or design? And a protocol or set of rules? What was the study design or method: observational, experimental, case-control, prospective, retrospective, or what? (See the previous chapter for kinds of studies and their uses and limits.) "A lot of people just scrounge around and try to come up with some conclusion without any real plan or design at the start," one medical editor reports. *Was the design drawn before you started your study? What specific questions or hypotheses did you set out to test or answer?*

2503012518

Why did you do it that way? Do you think it was the right kind of study to get the answer to this question or problem?

Was it a true human experiment, if possible, with comparable groups picked at random for comparison? If not, why not? And what was the substitute?

If an investigator patiently — you hope — tells you about an acceptable-sounding design, that's worth a brownie point. If the answer is "Huh?" or a nasty one, that may tell you something else.

Are you presenting preliminary data or something fairly conclusive? Are you presenting a conclusion or a hypothesis for further study? "Preliminary" and "interesting" can mean "unproved."

If the result is not reasonably conclusive, should there be further studies and what kind?

How many subjects, patients, cases, or people are you talking about? Are these numbers large enough, statistically rigorous enough, to get the answers you want? Was there an adequate number of patients to show a difference between treatments? Why are you calling a press conference to report on four patients?

Small numbers can sometimes carry weight. And they may sometimes be the only ones possible. "Sometimes small samples are the best we can do," one researcher says. But larger numbers are always more likely to pass statistical muster.

The number studied can also depend on the subject. A thorough physiological study of five cases of some difficult disorder may be important. One new case of smallpox would be a shocker in a world in which smallpox has supposedly been eliminated. In June 1981 the federal Centers for Disease Control reported that five young men, all active homosexuals, had been treated for *Pneumocystis carinii* pneumonia at three Los Angeles hospitals.¹ This alerted the world to what soon became the AIDS epidemic.

Who were your subjects? How were they selected? What were your criteria for admission to the study? Were rigorous laboratory tests used to

define the patients, or

Was the assignment randomized? Random chance of being assigned to the study best. How was the randomization done?

If the subjects i
"If it is a nonrar
some extraordina

Was there a c
always be weake:
son? In other wc
what are you comy
control group simil.
studied?

Vogt calls *
bly . . .
ular litera

*Do you have
ative of the genera
the disease or con
long way towa
are the results of.*

If your group
important population
statistical adjustment
specific groups, or
people, to make
nearly comparable
and stra

*Was the st
treatment with a*

2503012519

QUESTIONS REPORTERS CAN ASK

51

define the patients, or were clinical diagnoses (necessarily less reliable) used? Was the assignment of subjects to treatment or other intervention randomized? Randomization should give every patient a 50 percent chance of being assigned to one group or the other of a two-armed study (one comparing two groups). Were the patients admitted to the study before the randomization? This helps eliminate bias.

How was the randomization done?

If the subjects weren't randomized, why not? One statistician says, "If it is a nonrandomized study, a biased investigator can get some extraordinary results by carefully picking his subjects."

Was there a control or comparison group? If not, the study will always be weaker. Who or what were your controls or bases for comparison? In other words: When you say you have such and such a result, what are you comparing it with? Are the study or patient group and the control group similar in all respects but the treatment or other variable being studied?

Vogt calls "comparison of non-comparable groups probably . . . the single most common error in the medical and popular literature on health and disease."¹⁷

Do you have reason to believe your subjects and controls were representative of the general population? Or the particular population — those with the disease or condition you are interested in? The answers here go a long way toward answering these questions: To what populations are the results applicable? Would the association hold for other groups?

If your groups are not comparable to the general population or some important populations, have you taken steps to adjust for this? Either statistical adjustment or stratification of your sample to find out about specific groups, or both? Samples can be adjusted for age, for example, to make an older- or younger-than-average sample more nearly comparable to the general populace. (More on applicability and stratification after a bit.)

Was the study blind? In a study comparing drugs or other forms of treatment with a placebo or dummy treatment, did (1) those administering

2503012520

the treatment, (2) those getting it, and (3) those assessing the outcome know who was getting what, or were they indeed blinded, knowing only that they were comparing A and B (or A, B, and C, perhaps)?

Could those giving or getting the treatment have easily guessed which was which by a difference in reaction or taste or other results?

Not every study can be a blind study. One researcher says, "There can be ethical problems in not telling patients what drug they're taking and the possible side effects. People are not guinea pigs." True enough, but a blinded study will always carry more conviction.

Were there other accepted quality controls? For example, making sure (perhaps by counting pills or studying urine samples) that the patients supposed to take a pill really took it.

Were you able to follow your protocol or study plan?

If there were questionnaires, interviews, or a survey: *Were the questions likely to elicit accurate, reliable answers? Was it really possible to get accurate answers to these questions?*

Sampling is as common in medical studies as in political polling. Every study examines a sample, not the whole population. The sample must be reasonably accurate to give valid results. But badly worded questions can also distort the results. Respondents' answers can differ sharply, depending on how questions are asked. Example: In one study 1,153 subjects were asked which is safer, a treatment that kills 10 percent of every 100 patients or a treatment with a 90 percent survival rate? More people voted for the second way of saying precisely the same thing.³

People commonly give inaccurate answers to sensitive questions, such as those about sexual behavior. They are notoriously inaccurate in reporting their own medical histories, even those of recent months.

Ask: *Did you pretest your questions for effectiveness before doing your actual survey?*

Also: *What was your nonresponse rate? Do you report it?*

In any study course? Do you accou

Every study David Sackett say reasons. Rather, recover, die, or reability." If an inve dropped out, it ca died of "other ca being investigated after all, they di treatment look b deaths in every t

Sackett add: original inception more are not ac worth reading."

"Generally true, Professor W few relat contain all r sometimes been .. including or what extent has data?" . . It is data to make t

How long u
ily survive with i
really know the c

And: Hou
bias—a dis
made by findi
but a cure wa
"It does pay to

2503012521

In any study: *How many of your study subjects completed the course? Do you account for those who dropped out and tell why they did?*

Every study has dropouts. McMaster University's Dr. David Sackett says, "Patients do not disappear . . . for trivial reasons. Rather, they leave . . . because they refuse therapy, recover, die, or retire to the Sunbelt with their permanent disability." If an investigator ignores those who didn't do well and dropped out, it can make the outcome look better. If those who died of "other causes" are listed among "survivors" of the disease being investigated — this is sometimes done on the theory that, after all, they didn't die of the target cause — it can make a treatment look better unless there are equal numbers of such deaths in every branch of the study.

Sackett adds, "The loss to follow-up of 10 per cent of the original inception cohort is cause for concern. If 20 per cent or more are not accounted for, the results . . . are probably not worth reading." (On which Dr. Thomas Vogt comments, "Generally true, but utterly dependent on the situation.")

Professor Warren Burkett of the University of Texas adds a few related and pointed questions: "Does the paper or publication contain all results of all experiments? Support for a hypothesis has sometimes been made to seem stronger by selective reporting . . . including only the data that most closely fit the theory. To what extent has the data offered . . . been smoothed from the raw data?" . . . It is not unknown for researchers to clip and round data to make them fit [their] predicted results" (italics mine).⁵

How long was the study's follow-up? How long do patients ordinarily survive with this disease? Were your patients followed long enough to really know the outcomes, good or bad?

And: *How thorough was the follow-up?* In one report on amoebiasis — a disease caused by an amoeba — the diagnosis was made by finding the amoeba in one of three consecutive stools, but a cure was declared after observing just one negative stool. "It does pay to read with care," a medical professor observes.

2503012522

Could your results have occurred just by chance? Have any statistical tests been applied to test this?

Did you calculate a P value? Was it favorable—.05 or less? (Reported as <.05; see Chapter 3.) P values and confidence statements need not be regarded as straitjackets, but like jury verdicts, they indicate reasonable doubt or reasonable certainty.

Remember that positive findings are more likely to be reported and published than negative findings. Remember that a favorable-sounding *P* value of <.05 means only that there is just 1 chance in 20, or a 5 percent probability, that the statistics could have come out this way by pure chance *when there was actually no effect*—so 1 in every 20 statistically significant results may be a misleading false positive.

There are also ways and ways of arriving at *P* values. For example, an investigator may choose to report one of several *end points*: death, length of survival, blood pressure, other measurements, or just the patient's condition on leaving the hospital. All can be important, but a *P* value can be misleading if the wrong one is picked or emphasized.

You might want to ask: *Are all the important end points and their P values reported? Also: Was the test giving the P value the appropriate test, as planned in your written protocol, or did you finally do more than one kind of test? (And perhaps report only the best answer?) What were the other values?*

Did you collaborate with a statistician in both your design and your analysis? A statistician's collaboration often may be indicated in a credit or footnote.

In studies seeking *cause and effect*, remember that association is not necessarily causation. Rutgers' Dr. Michael Greenberg reminds us, "Mathematical methods cannot establish proof of cause and effect. They can indicate the probability that a relationship occurred by chance, can sometimes quantify the existing relationship between actions and effects, and can under the best circumstances be used to predict the impact of actions even

if the comple
... View m
skepticism."

A true e
prove cause
and chemistr
association i:
experiment)
ria that you

*Is the ass
different pla*

*How str.
describing a
ratio? The w
It mainly m
ing the outc*

*A relati
one by the c
(see !
55 to*

*188 pe. 10
smokers wi
cancer—the*

*Is there
curve or gr
agent, or c
deed at gr
smokers at
is an unsett
only after :*

*Anoth
correlation co
the associa
tion, betwe
straight, st
a straight*

2503012523

if the complex phenomena driving them are not understood. . . . View mathematical associations with a healthy degree of skepticism."

A true experiment, controlling all variables, can sometimes prove cause and effect almost surely. This is easier in physics and chemistry than in human biology. When, then, does a close association in an observational study (rather than a controlled experiment) indicate causation? There are several possible criteria that you can ask about:

Is the association consistent? Are similar results usually found in different places and by different research methods?

How strong is the association? If risk is an appropriate way of describing a particular situation: *What is the relative risk, or the risk ratio?* The word "strong" is used here in its mathematical sense. It mainly means the *magnitude* of an effect or risk, the *odds* favoring the *outcome of interest* versus no such outcome.

A *relative risk*, or *risk ratio*, compares two rates by dividing one by the other. In an American Cancer Society smoking study (see page 46) the lung cancer mortality rate in nonsmokers aged 55 to 69 was 19 per 100,000 per year; the risk in smokers was 188 per 100,000. Since 188 divided by 19 equals 9.89, the smokers were about 9.9 times more likely to die from lung cancer—their relative risk was 9.9.⁶ That's strong!

Is there an impressive dose-response, or cause-and-effect, curve—a curve or gradient that shows that the greater the exposure to the agent, or cause, the greater the effect? Heavy smokers are indeed at greater risk than moderate smokers, and moderate smokers at greater risk than light smokers. (In some cases—this is an unsettled matter—there may be a threshold effect, an effect only after some minimum dose.)

Another way of asking about risk and response: *What is the correlation coefficient*—the extent to which a set of measurements of the association is *linear*? A perfect linear relationship, or correlation, between two observations or variables would show up as a straight, steadily rising set of *data points*—in everyday language, a straight line on a graph. A perfect positive correlation or,

2503012524

linear relationship, is given the value +1; +.5 would be a lesser but still interesting relationship; -1 or any negative figure indicates an *inverse* or *negative relationship*, such as a runner's speed going down as his weight goes up. A correlation of zero means no consistent association.

How specific is the association? Does a supposed cause lead to many supposed effects? Or does an effect depend on many supposed causes? Such associations are less specific, and thus more suspect, until positive evidence piles up. Smoking indeed causes many effects. A lung disease, asbestosis, is most common when there is exposure to both asbestos and cigarette smoke.

Does the supposed cause precede the effect? Is a supposed biological association epidemiologically plausible? One strong argument for a cause-and-effect relationship between high consumption of saturated fats and cholesterol and coronary heart disease is that populations on such diets generally develop more such disease than those on leaner diets.

Does the association make biological sense? Does it agree with current biological and physiological knowledge? You can't follow this test out the window. Much biological fact is ill understood. Also, Mosteller warns, "Someone nearly always will claim to see a [biological or physiological] association. But the people who know the most may not be willing to."

Finally, look for the real why. Ask: *Are there other possible explanations? Did you look for other explanations—confounders, or confounding variables*, that may be producing or helping produce the association? Sometimes we read that married people live longer than singles. Does marriage really increase life span, or may medical or other problems make some people less likely to marry and also die sooner? Maybe the Dutch thought storks brought babies because better-off families had more chimneys, more storks, and more babies.

Did you take steps to control or adjust for other possible explanations? Did you do a stratified analysis—a breakdown of the data by strata like sex, race, socioeconomic status, geographical area, occupation? Men commonly have more bronchitis and cirrhosis of the

liver than w
more heart
possibly beca
analyses will

Did you
ate analysis) &
analyses can
also be misu
Some sophis
analyses did y
the more an
consider? How
tor tries eno
tion, he or
untrue.

In caus
reanalysis of
independen
see if

or reanalyse
analysis or i
among auth
reasoned ar
than the an

In stud
know or deci
explicit, objec
ments or te
terviews, pt
highly subj
provement
quantify) o
Was there so
If two c

2503012525

QUESTIONS REPORTERS CAN ASK

57

liver than women because they drink more. They also have more heart disease, possibly because they've smoked longer, possibly because some hormones protect women. Only stratified analyses will bring out such differences.

Did you do an analysis (a regression or some other form of multivariate analysis) to try to identify the important variable or variables? Such analyses can often reveal the strongest associations. They can also be misused, and they are not always needed or appropriate. Some sophisticated questions, when appropriate: How many such analyses did you have to run to decide on the appropriate one? Sometimes the more analyses, the worse the study. How many variables did you consider? How many of these did you wind up reporting? If an investigator tries enough variables in a kind of statistical fishing expedition, he or she is almost bound to find something, true or untrue.

In cause-and-effect and other studies, ask: *Has there been any reanalysis of the data? "Results, if possible, should be method-independent," Greenberg believes. "You should recalculate and see if the results hold up."*

A word of caution: Questions about multivariate analyses or reanalyses can be tricky. Whether or not to do one kind of analysis or reanalysis or none at all is often a matter of dispute among authorities. Launch the subject with some humility. A reasoned answer, affirmative or negative, may tell you more than the answer's precise content.

In studies of medical treatments or preventives: *How did you know or decide when your patients were cured or improved? Were there explicit, objective outcome criteria? That is, were there firm measurements or test results rather than physicians' observations in interviews, physical examinations, or chart reviews, all techniques highly subject to great observer variation and inaccuracy? If improvement or relief from pain — a particularly soft (hard to quantify) outcome measure — had to be judged by observers:*

Was there some systematic way of making an assessment? If two or more groups were compared for survival, was their starting

2503012526

point the same at onset? At diagnosis? At start of treatment? Were they judged by the same disease definitions at the start and the same measures of severity and outcome?

Did the intervention have the good results that were intended? Has there been an evaluation to see whether it was a useful result?

Investigators often report that a drug or other measure has lowered blood cholesterol levels. Fine, but were they able to show that it reduced the number of heart attacks? Or was reduction of a supposed risk factor itself taken to mean the hoped-for outcome? That may often be necessary, but the issue should be discussed.

Investigators once reported that a new heart drug reduced the number of recurrent myocardial infarctions (heart attacks), fatal and nonfatal. But total mortality for all causes was higher in the treated group than in a placebo group.

Public health officials may announce the success of a campaign to take high blood pressure measurements: X number of people were found to be hypertensive and were referred to their doctors. But how many went to their doctors? How many of those received optimum treatment? Were their blood pressures reduced? (If they were, the evidence is strong that they should suffer fewer strokes.)

In short: *What was the bottom line? Did you really do any good?*

To whom do your results apply? Can they be generalized to a larger population? Are your patients like the average doctor's patients? Is there any basis in these findings for any patient to ask his or her doctor for a change in treatment? Clinic populations, hospital populations, and the "worst cases" are not necessarily typical of patients in general, and improper generalization is unfortunately common in the medical literature.

Again and again, in many of the cases cited in this chapter, ask: Do other studies back you up? Are your results consistent with other clinical and experimental findings? Have your results been repeated or

*confirmed or su
these results?*

Virtually
studies add c
criteria and t
in humans, :

One sci
grab bag of
circumstances.
but consider
John Bailar
several low
integrating
than any or

Mostly
most impor
these: *What*
data really
lated b- or
mad

Lives at
and flaws in
the investigat
Robert Be
audacity a:
use qualifi
bound to i

Ask t
your work
rienced sc
ers gener:

—
"Freder
common sen
thought of it

2503012527

confirmed or supported by other studies? Or have only you been able to get these results?

Virtually no single study proves anything. Two or 4 or 15 studies add credence, especially if the diagnostic and outcome criteria and the people studied are similar. Consistency of results in humans, animals, and laboratory tests also adds credence.

One scientist warns, however, "You have to be wary about a grab bag of studies with different populations and different circumstances." To which Harvard's Mosteller adds, "Yes, be wary, but consistency across such differences cheers me up." And Dr. John Bailar tells us that, despite possible pitfalls, "meta-analysis of several low power reports"—that is, statistically analyzing and integrating their results—"may come to stronger conclusions than any one of them alone" (italics mine).⁸

Mostly just good-sense questions? Of course. Some of the most important questions of all for a reporter to ponder are these: *What do I think? Do the conclusions make sense to me?* Do the data really justify the conclusions? If this person has extrapolated beyond the evidence, has he or she explained why and made sense?⁹

Does the investigator frankly document or discuss the possible biases and flaws in the study? A good scientific paper should do so. *Does the investigator admit that the conclusion may be tentative or equivocal?* Dr. Robert Boruch of Northwestern University says, "It requires audacity and some courage to say, 'I don't know.'" *Do the authors use qualifying phrases?* If such phrases are important, we are bound to include them in any responsible story.

Ask the investigators themselves: *How much weight should your work be given? Is it really firm? And how important?* An experienced science reporter says, "I have found that good researchers generally have an honest and proportionate view of their

⁸Fredrick Mosteller disagrees with my occasional reference to good sense or common sense. If something is a commonsense idea, he says, "Surely all would have thought of it. So it must be uncommon sense after all." He makes good sense.

own work's importance." But there are many exceptions.

Ask others in the same field: *How do other informed people regard this report—and these investigators? Are they speaking in their own area of expertise, or have they shown real mastery if they have ventured outside it? Have their past results generally held up? And what are some good questions I can ask them?* True, a lot of brilliant and original work has been pooh-poohed for a time by others. Still, scientists survive only by eventually convincing their colleagues.

More formally: *Has there been a review of the data and conclusions by any disinterested parties?* Some major clinical studies are reviewed by independent second parties or committees. Reports of the National Academy of Sciences must pass muster by a review committee.

Has there been peer review of the material? That is, has it been examined by referees who were sent the article by a journal editor?

And, a very important question: *Has the work been published or accepted by a reputable journal? If not, why not?* The *New England Journal of Medicine* prints only 15 percent of the papers submitted to it (many, of course, are rejected because they are not of enough interest to the journal's readers). Many have been given at medical or scientific meetings, yet do not pass peer reviewers' or the editors' muster. Most are eventually published elsewhere, many in good journals. But there are journals and journals.

In science as a whole, including biology and often basic medical sciences, *Science* and the *British Nature* are indispensable. In general medicine and clinical science at the physician's level, the best, most useful journals are probably *New England Journal of Medicine*, *Journal of the American Medical Association*, *Annals of Internal Medicine*, *Canadian Medical Journal*, *Journal of Clinical Investigation*, and the British *Lancet* and *British Medical Journal*. There are many equally good specialty journals as well as mediocre ones. In epidemiology, three good sources are *American Journal of Epidemiology*, *Journal of Chronic Diseases*, and *Preventive Medicine*. Ask people in any field: What are the most reliable journals, those where you would want your work published?

Some of
are not journ:
like *Family Pr*
primary articles:
free-circulatio
and medical r
revenue, are
journals. Th
journals prin
ords of wor
Journal's Dr.

Read th
the investiga
the article h
library, whic
hospitals, ar
cieties. Too :
conservative
further in ir
to go in the
review " u
put ;
read the

Most r
an article, b
• A cri
tician, and

ysis and its
to detect tr
at least ass
statistical a
times. Son
isn't identif
tics.

• Tabl
sions. Sor

2503012529

QUESTIONS REPORTERS CAN ASK

61

Some of the most valuable journals to a medical reporter are not journals of original publication but review publications like *Family Practice* and *Hospital Practice*, which mainly print summaries articles for practitioners. With some strong exceptions, the free-circulation — also known as controlled-circulation — journals and medical magazines, which depend wholly on advertising for revenue, are not as rigorously screened as the traditional journals. They are often on top of the news, however. All journals print clinkers sometimes. "Scientific journals are records of work, not of revealed truth," says the *New England Journal's* Dr. Arnold Reisman.¹⁰

Read the entire journal article yourself, if there is one. Ask the investigator for a copy or phone the journal. Or, assuming the article has already been published, look for it at a medical library, which can be found at any medical college, most good hospitals, and the headquarters of many county medical societies. Too many news releases tout articles that read far more conservatively than the PR version. Many scientists go much further in interviews or news conferences than they are willing to go in their articles. A reporter asked a scientist, "Does peer review of an article put you at ease?" He said, "It should help put you at greater ease, but nothing puts me at ease until I've read the article."

Most reporters can't be scientific referees, but when you read an article, look for the following:

- A credit or footnote indicating collaboration with a statistician, and a paragraph describing the method of statistical analysis and its outcomes, such as *P* value or confidence level, power to detect treatment effects, and so on. If they're in place, you can at least assume that some effort was made to apply the rigors of statistical analysis. If they're missing, should you beware? Sometimes. Sometimes the statistician is a coauthor whose specialty isn't identified. And some investigators are well versed in statistics.
- Tables and figures that tell the same story as the conclusions. Sometimes they don't. One statistician told reporters,

2503012530

"Don't assume that someone can interpret his own data. You may do better." And "muddle around in the footnotes and appendices," Mosteller advises. "You might find a few horrors. That's how people found out that a much publicized study of public and private schools included only about 12 private, non-parochial schools."

• Other things described in this chapter, such as the protocol and study design, the criteria for admitting and randomizing subjects, the therapy actually received (in contrast to that planned in the protocol), blinding, complications, loss to follow-up, follow-up time, and any discussion of reservations or weaknesses.

Ask, when appropriate: *Where did the money to support the study come from?*² Many honest investigators are financed by companies that may profit from the outcome. So are some dishonest or self-deluding investigators. But the peddler of a biased point of view is as likely to be an antiestablishment crusader—or an academic ladder-climber—as a corporate darling. Perhaps the best question to ask yourself is, Is this investigator a scientist or a salesman? In any case, the public should know any pertinent connections.

"What proportion of papers will satisfy [all] the requirements for scientific proof and clinical applicability?" Sackett writes, "Not very many. . . . After all, there are only a handful of ways to do a study properly but a thousand ways to do it wrong."¹¹

Despite impeccable design, some studies yield answers that turn out to be wrong. Some fail for lack of understanding of physiology and disease. Even the soundest studies may provoke controversy. No study settles anything for all time.

And according to Sackett, some "may meet considerable resistance when they discredit the only treatment currently available. . . . Clinicians may still elect to do something, even if it is of no demonstrable benefit. Study results may be rejected,

regardless of
hood of the
Report
everything
some of th-

2503012531

QUESTIONS REPORTERS CAN ASK

63

regardless of their merit, if they threaten the prestige or livelihood of their audience."

Reporters need to tread a narrow path between believing everything and believing nothing. Also—we are reporters—some of the controversies make important stories.

2503012532

6

Tests and Testing

Testing is often the only way to answer our questions, but it doesn't produce unassailable, universal truths that should be carved on stone tablets. Instead, testing produces statistics, which must be interpreted.

— Robert Hooke

Who knows when thou mayest be tested?

— Ronald Arthur Hopwood

DO physicians always know what they're doing when they administer tests? Stanford's Dr. Eugene Robin says many tests "have not been properly evaluated and in fact may be useless or harmful." He asks, "Is it common practice in medicine to perform careful clinical trials before introducing tests that can affect the welfare of masses of patients? Sadly, the answer is no."

A good test should detect both health and disease and do so with high accuracy. The measures of the value of a *clinical test*, one used for medical diagnosis, are *sensitivity* and *specificity*, or, simply, the ability to avoid *false negatives* and *false positives*. *Sensitivity* is how well a test identifies a disease or condition in those who have it—how well it *avoids false negatives*, or missed cases. If 100 people with a condition are tested and 90 test positive, the test's sensitivity is 90 percent. *Specificity* is how well a test identifies those who do not have the disease or condition—how well it *rules out false positives*, or mistaken identifications. If 100 healthy people are tested and 90 test negative, the test's specificity is 90 percent.

Sensitivity, in short, tells us about *disease present*. *Specificity* tells us about *disease absent*. A highly unspecific test will produce many false positives; a highly insensitive test, many false nega-

64

tives. Almost qualities—such an overlap. In every case, though you will get labeling, the you will get.

As a border terms. ("So comments.) concept, the fact that test person who this:

How many know this? How tests in the medical context some follow-up, condition be subjects, are frequent.

How many false positives not to miss sensitivity to avoiding failure anyway, on

Doubt because in acceptable short, then ruled out from detected pr

2503012533